

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.
OM protein - protein search, using sw model
Run on: October 8, 2004, 17:09:12 ; Search time 136 Seconds
(without alignments)
1616.094 Million cell updates/sec

Title: US-09-940-921B-2
Perfect score: 3616
Sequence: 1 MDKYDVIRKAIQGGAFKAYL.....SVSRQEGTKDPYSPVLILM 683

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1351062 seqs, 321799191 residues

Total number of hits satisfying chosen parameters: 1351062

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications AA:*
1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodata/2/pubpaa/PTCT_NEW_PUB.pep:*
3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep:*
4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep:*
5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep:*
6: /cgn2_6/ptodata/2/pubpaa/PTCT_PUBCOMB.pep:*
7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep:*
8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep:*
9: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
10: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
11: /cgn2_6/ptodata/2/pubpaa/US09_PUBCOMB.pep:*
12: /cgn2_6/ptodata/2/pubpaa/US09_NEW_PUB.pep:*
13: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
14: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
15: /cgn2_6/ptodata/2/pubpaa/US10_PUBCOMB.pep:*
16: /cgn2_6/ptodata/2/pubpaa/US10_NEW_PUB.pep:*
17: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep:*
18: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3616	100.0	683	9 US-09-940-921B-2	Sequence 2, Appli
2	3431.5	94.9	654	9 US-09-940-921B-4	Sequence 4, Appli
3	3081.5	85.2	640	14 US-10-243-735-2	Sequence 2, Appli
4	3081.5	85.2	640	16 US-10-730-010-2	Sequence 2, Appli
5	2111.5	58.4	403	12 US-10-114-270-188	Sequence 186, App
6	2106.5	58.3	399	12 US-10-114-270-186	Sequence 186, App
7	1106.5	30.6	774	14 US-10-162-706-5	Sequence 5, Appli
8	1081.5	29.9	1214	9 US-09-783-320-4	Sequence 4, Appli
9	1066	29.5	616	14 US-10-243-735-4	Sequence 4, Appli
10	1065	29.5	616	16 US-10-730-010-4	Sequence 4, Appli
11	926.5	25.6	254	10 US-09-898-837A-35	Sequence 35, Appl
12	818	22.6	506	12 US-10-114-270-180	Sequence 180, App
13	816.5	22.6	345	9 US-09-870-962-6	Sequence 6, Appli
14	812	22.5	425	12 US-10-114-270-184	Sequence 184, App
15	812	22.5	459	12 US-10-114-270-182	Sequence 182, App

16	806	22.3	506	12	US-10-415-011-7	Sequence 7, Appli
17	728.5	20.1	459	9	US-09-771-161A-198	Sequence 198, App
18	683	18.9	850	12	US-10-425-114-56926	Sequence 56926, A
19	683	18.9	850	12	US-10-425-114-56928	Sequence 56928, A
20	668	18.5	841	14	US-10-198-070-52	Sequence 52, Appl
21	668	18.5	841	16	US-10-408-765A-1685	Sequence 1685, Ap
22	668	18.5	841	16	US-10-620-052A-32	Sequence 32, Appl
23	660	18.3	841	14	US-10-198-070-87	Sequence 87, Appl
24	645	17.8	416	9	US-09-731-231A-6	Sequence 6, Appli
25	645	17.8	416	16	US-10-751-985-6	Sequence 6, Appli
26	643	17.8	256	16	US-10-620-052A-48	Sequence 48, Appl
27	638	17.6	260	9	US-09-731-231A-5	Sequence 5, Appli
28	638	17.6	260	16	US-10-751-985-5	Sequence 5, Appli
29	634.5	17.5	713	14	US-10-162-706-2	Sequence 2, Appli
30	634	17.5	692	9	US-09-932-481-2	Sequence 2, Appli
31	634	17.5	692	14	US-10-434-034-2	Sequence 2, Appli
32	633	17.5	698	15	US-10-074-978A-144	Sequence 144, App
33	632.5	17.5	280	14	US-10-162-706-6	Sequence 6, Appli
34	632.5	17.5	460	12	US-10-362-892-23	Sequence 23, Appl
35	632.5	17.5	460	14	US-10-162-706-4	Sequence 4, Appli
36	632.5	17.5	460	15	US-10-288-798-23	Sequence 23, Appl
37	632	17.5	255	10	US-09-898-837A-36	Sequence 36, Appl
38	632	17.5	692	9	US-09-910-150-5	Sequence 5, Appli
39	632	17.5	692	12	US-10-377-097-5	Sequence 5, Appli
40	627.5	17.4	326	15	US-10-074-978A-16	Sequence 16, Appl
41	623	17.2	696	15	US-10-074-978A-146	Sequence 146, App
42	617.5	17.1	291	14	US-10-162-706-17	Sequence 17, Appl
43	617.5	17.1	291	15	US-10-074-978A-145	Sequence 145, App
44	615.5	17.0	253	15	US-10-074-978A-149	Sequence 149, App
45	602	16.6	645	9	US-09-731-231A-2	Sequence 2, Appli

ALIGNMENTS

RESULT 1

US-09-940-921B-2
; Sequence 2, Application US/09940921B
; Patent No. US20020147320A1
; GENERAL INFORMATION:
; APPLICANT: Friddle, Carl Johan
; APPLICANT: Hilbun, Erin
; APPLICANT: Nepomnichy, Boris
; APPLICANT: Hu, Yi
; TITLE OF INVENTION: No. US20020147320alel Human Kinase Proteins and Polynucleotides E
; FILE REFERENCE: LEX-0227-05A
; CURRENT APPLICATION NUMBER: US/09/940.921B
; CURRENT FILING DATE: 2002-05-21
; PRIOR APPLICATION NUMBER: US 60/229,280
; PRIOR FILING DATE: 2000-08-31
; NUMBER OF SEQ ID NOS: 10
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 683
; TYPE: PRT
; ORGANISM: homo sapiens
US-09-940-921B-2

Query Match	100.0%;	Score 3616;	DB 9;	Length 683;
Best Local Similarity	100.0%;	Pred. No. 8.6e-245;		
Matches 683;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	MDKYDVIRKAIQGGAFKAYLAKGKSDSKHCVIKINFEEKMPIQKEASKKEVILEKMKH	60	
Db	1	MDKYDVIRKAIQGGAFKAYLAKGKSDSKHCVIKINFEEKMPIQKEASKKEVILEKMKH	60	
QY	61	PNIVAFNFSQENGRFLTWMEYCDGDLMKRINFORGVLFSEDIILGFVQISLGLKHII	120	
Db	61	PNIVAFNFSQENGRFLTWMEYCDGDLMKRINFORGVLFSEDIILGFVQISLGLKHII	120	
QY	121	DRKILHRDIKQNIFLSKNGWVAKLGDGFIARVLNNSMELARTICIGTPYILSPICQNKP	180	
Db	121	DRKILHRDIKQNIFLSKNGWVAKLGDGFIARVLNNSMELARTICIGTPYILSPICQNKP	180	

QY 181 YNNKTDIWSLGCVLVELCTLKHPFEGNNLQOLVLKICQAHFAPISPGFSREHLSLSQLF 240
 DB 181 YNNKTDIWSLGCVLVELCTLKHPFEGNNLQOLVLKICQAHFAPISPGFSREHLSLSQLF 240
 QY 241 QVSPDRPSINSILKRPFLNLPKYLTPQENTGVEDYQOETRHGPPSPQWPAEYLQKFEAQYK 300
 DB 241 QVSPDRPSINSILKRPFLNLPKYLTPQENTGVEDYQOETRHGPPSPQWPAEYLQKFEAQYK 300
 QY 301 IQKVRFGCKPPRSRISVPIKRNAILHNRNWRPPAGAQAQKARSIKMTERPKIAAACHYDY 360
 DB 301 IQKVRFGCKPPRSRISVPIKRNAILHNRNWRPPAGAQAQKARSIKMTERPKIAAACHYDY 360
 QY 361 YVQALDMLRRRAHKKPSYHPIQENTGVEDYQOETRHGPPSPQWPAEYLQKFEAQYK 420
 DB 361 YVQALDMLRRRAHKKPSYHPIQENTGVEDYQOETRHGPPSPQWPAEYLQKFEAQYK 420
 QY 421 VEKQGLRPSAPNNQOELRSNGEPRFQELPFRKNEMKEQYWKOLEIRQOYHND 480
 DB 421 VEKQGLRPSAPNNQOELRSNGEPRFQELPFRKNEMKEQYWKOLEIRQOYHND 480
 QY 481 MKEIRKMGREPEENSKISHKTYLVKKSNNLPVHQDASEGEAPVQDIEKDKQMLQNTKE 540
 DB 481 MKEIRKMGREPEENSKISHKTYLVKKSNNLPVHQDASEGEAPVQDIEKDKQMLQNTKE 540
 QY 541 SKNPEQKYKAKGVKFEINLDCISDENILQEEAMDIPTNETLTFEDGMKPKFEYECVKEH 600
 DB 541 SKNPEQKYKAKGVKFEINLDCISDENILQEEAMDIPTNETLTFEDGMKPKFEYECVKEH 600
 QY 601 GYTDKAFELKHCPEAGFSTQTVAAVGNRRQWDGGAPQTLQWMAVADITSTCPTGPDSE 660
 DB 601 GYTDKAFELKHCPEAGFSTQTVAAVGNRRQWDGGAPQTLQWMAVADITSTCPTGPDSE 660
 QY 661 SVLSVSRQEGKTKDPYSPVLILM 683
 DB 661 SVLSVSRQEGKTKDPYSPVLILM 683

RESULT 2
 US-09-940-921B-4
 ; Sequence 4, Application US/09940921B
 ; Patent No. US20020147320A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Fiddle, Carl Johan
 ; APPLICANT: Hilbun, Erin
 ; APPLICANT: Nepomnichy, Boris
 ; APPLICANT: Hu, Yi
 ; TITLE OF INVENTION: No. US20020147320A1 Human Kinase Proteins and Polynucleotides
 ; FILE REFERENCE: LEX-0227-USA
 ; CURRENT APPLICATION NUMBER: US/09/940,921B
 ; CURRENT FILING DATE: 2002-05-21
 ; PRIOR APPLICATION NUMBER: US 60/229,280
 ; PRIOR FILING DATE: 2000-08-31
 ; NUMBER OF SEQ ID NOS: 10
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 4
 ; LENGTH: 654
 ; TYPE: PRT
 ; ORGANISM: homo sapiens
 US-09-940-921B-4

Query Match 94.9%; Score 3431.5; DB 9; Length 654;
 Best Local Similarity 95.8%; Pred. No. 6.9e-232;
 Matches 654; Conservative 0; Mismatches 0; Indels 29; Gaps 1;

QY 1 MDKYDVIKAIQOGAFKAYLAKGSDSKHCVTKEINFEKMPIQEKESKKEVILEKMKH 60
 DB 1 MDKYDVIKAIQOGAFKAYLAKGSDSKHCVTKEINFEKMPIQEKESKKEVILEKMKH 60
 QY 61 PNIVAFNFSQNGRLFTVMEYCDGDLMKRINRQGVLFSEDDQILGWVQISLGLKH 120
 DB 61 PNIVAFNFSQNGRLFTVMEYCDGDLMKRINRQGVLFSEDDQILGWVQISLGLKH 120

QY 121 DRKTLHRDIIKAQNIPLSKNGWAKLGDGFIARVLNNMELARTCIGTPYVLSPEICQNK 180
 DB 121 DRKTLHRDIIKAQNIPLSKNGWAKLGDGFIARVLNNMELARTCIGTPYVLSPEICQNK 180
 QY 181 YNNKTDIWSLGCVLVELCTLKHPFEGNNLQOLVLKICQAHFAPISPGFSREHLSLSQLF 240
 DB 181 YNNKTDIWSLGCVLVELCTLKHPFEGNNLQOLVLKICQAHFAPISPGFSREHLSLSQLF 240
 QY 241 QVSPDRPSINSILKRPFLNLPKYLTPQENTGVEDYQOETRHGPPSPQWPAEYLQKFEAQYK 300
 DB 241 QVSPDRPSINSILKRPFLNLPKYLTPQENTGVEDYQOETRHGPPSPQWPAEYLQKFEAQYK 300
 QY 301 IQKVRFGCKPPRSRISVPIKRNAILHNRNWRPPAGAQAQKARSIKMTERPKIAAACHYDY 360
 DB 301 IQKVRFGCKPPRSRISVPIKRNAILHNRNWRPPAGAQAQKARSIKMTERPKIAAACHYDY 360
 QY 361 YVQALDMLRRRAHKKPSYHPIQENTGVEDYQOETRHGPPSPQWPAEYLQKFEAQYK 420
 DB 361 YVQALDMLRRRAHKKPSYHPIQENTGVEDYQOETRHGPPSPQWPAEYLQKFEAQYK 420
 QY 421 VEKQGLRPSAPNNQOELRSNGEPRFQELPFRKNEMKEQYWKOLEIRQOYHND 480
 DB 421 VEKQGLRPSAPNNQOELRSNGEPRFQELPFRKNEMKEQYWKOLEIRQOYHND 480
 QY 481 MKEIRKMGREPEENSKISHKTYLVKKSNNLPVHQDASEGEAPVQDIEKDKQMLQNTKE 540
 DB 481 MKEIRKMGREPEENSKISHKTYLVKKSNNLPVHQDASEGEAPVQDIEKDKQMLQNTKE 540
 QY 541 SKNPEQKYKAKGVKFEINLDCISDENILQEEAMDIPTNETLTFEDGMKPKFEYECVKEH 600
 DB 541 SKNPEQKYKAKGVKFEINLDCISDENILQEEAMDIPTNETLTFEDGMKPKFEYECVKEH 600
 QY 601 GYTDKAFELKHCPEAGFSTQTVAAVGNRRQWDGGAPQTLQWMAVADITSTCPTGPDSE 660
 DB 601 GYTDKAFELKHCPEAGFSTQTVAAVGNRRQWDGGAPQTLQWMAVADITSTCPTGPDSE 660
 QY 661 SVLSVSRQEGKTKDPYSPVLILM 683
 DB 661 SVLSVSRQEGKTKDPYSPVLILM 683

RESULT 3
 US-10-243-735-2
 ; Sequence 2, Application US/10243735
 ; Publication No. US20030022341A1
 ; GENERAL INFORMATION:
 ; APPLICANT: WEBSTER, Marion et al
 ; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
 ; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
 ; TITLE OF INVENTION: THEREOF
 ; FILE REFERENCE: CL001212DIV
 ; CURRENT APPLICATION NUMBER: US/10/243,735
 ; CURRENT FILING DATE: 2002-09-16
 ; NUMBER OF SEQ ID NOS: 4
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO 2
 ; LENGTH: 640
 ; TYPE: PRT
 ; ORGANISM: Human
 US-10-243-735-2

Query Match 85.2%; Score 3081.5; DB 14; Length 640;
 Best Local Similarity 94.1%; Pred. No. 2.3e-207;
 Matches 585; Conservative 2; Mismatches 4; Indels 31; Gaps 1;

QY 1 MDKYDVIKAIQOGAFKAYLAKGSDSKHCVTKEINFEKMPIQEKESKKEVILEKMKH 60
 DB 1 MDKYDVIKAIQOGAFKAYLAKGSDSKHCVTKEINFEKMPIQEKESKKEVILEKMKH 60
 QY 61 PNIVAFNFSQNGRLFTVMEYCDGDLMKRINRQGVLFSEDDQILGWVQISLGLKH 120
 DB 61 PNIVAFNFSQNGRLFTVMEYCDGDLMKRINRQGVLFSEDDQILGWVQISLGLKH 120

```
QY 121 DRKILHRDIKAQNI FLSKNGWAKLGD FGIARVLNNSMELARTICIGTPYLSPEICQK 180
DB 121 DRKILHRDIKAQNI FLSKNGWAKLGD FGIARVLNNSMELARTICIGTPYLSPEICQK 180
QY 181 YNNKTDIWSLGCVLVYELCTLKHPPFEGNNLQOLVLTICQAHFAPISPGFSRELHSLISOLF 240
DB 181 YNNKTDIWSLGCVLVYELCTLKHPPFEGNNLQOLVLTICQAHFAPISPGFSRELHSLISOLF 240
QY 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIQEEFSHMLICRAGAPASRHAGKVQCK 300
DB 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIQEEFSHMLICRAGAPASRHAGKVQCK 300
QY 301 IQKVRFGCKPPRSRSVPIKRNAILHRNEWRPPAGAKARSIKMIERPKIAAVCGHYD 360
DB 301 IQKVRFGCKPPRSRSVPIKRNAILHRNEWRPPAGAKARSIKMIERPKIAAVCGHYD 360
QY 361 YVQALDMLRRRAHKSYPHPIQENTGVEDYQGETRHGSPSQWPAEYLORKEFAQYK 420
DB 361 YVQALDMLRRRAHKSYPHPIQENTGVEDYQGETRHGSPSQWPAEYLORKEFAQYK 420
QY 421 VEKOLGLRPSAEPNPNQELRSNGEPRFQELPFRKNEMKEQEWKQLEIRQOYHND 480
DB 421 VEKOLGLRPSAEPNPNQELRSNGEPRFQELPFRKNEMKEQEWKQLEIRQOYHND 480
QY 481 MKEIRKMGREPEENSISKHTYLVKSNLPVHODASEGEAPVODIEKDLKQMLQNTKE 540
DB 481 MKEIRKMGREPEENSISKHTYLVKSNLPVHODASEGEAPVODIEKDLKQMLQNTKE 540
QY 541 SKNPEQKYKAKGVKFEINLDCISDENILQEEAMDI PNETLTFFDGMKFEYECVKEH 600
DB 541 SKNPEQKYKAKGVKFEINLDCISDENILQEEAMDI PNETLTFFDGMKFEYECVKEH 600
QY 601 GDYTDKAFKHLHCPAAGFSTQ 622
DB 570 GDYTDKAFKHLHCPAAGFSTQ 591
```

RESULT 4

```
US-10-730-010-2
; Sequence 2, Application US/10730010
; Publication No. US20040142366A1
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; THEREOF
; FILE REFERENCE: CLO01212CIPDIV-II
; CURRENT APPLICATION NUMBER: US/10/730,010
; CURRENT FILING DATE: 2003-12-09
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-730-010-2
```

```
Query Match 85.2%; Score 3081.5; DB 16; Length 640;
Best Local Similarity 94.1%; Pred. No. 2.3e-207;
Matches 585; Conservative 2; Mismatches 4; Indels 31; Gaps 1;
```

```
QY 1 MDKYDVIKAIQGAFGKAYLAKGSDSKHCVKIENFEXKPIQEKASKEVILLERKMH 60
DB 1 MDKYDVIKAIQGAFGKAYLAKGSDSKHCVKIENFEXKPIQEKASKEVILLERKMH 60
QY 61 PNIVAFNFSQENGRIFVMEYCDGDLMKRINRQGVLFSEDOILGWVQISLGHKHH 120
DB 61 PNIVAFNFSQENGRIFVMEYCDGDLMKRINRQGVLFSEDOILGWVQISLGHKHH 120
QY 121 DRKILHRDIKAQNI FLSKNGWAKLGD FGIARVLNNSMELARTICIGTPYLSPEICQK 180
DB 121 DRKILHRDIKAQNI FLSKNGWAKLGD FGIARVLNNSMELARTICIGTPYLSPEICQK 180
```

```
QY 181 YNNKTDIWSLGCVLVYELCTLKHPPFEGNNLQOLVLTICQAHFAPISPGFSRELHSLISOLF 240
DB 181 YNNKTDIWSLGCVLVYELCTLKHPPFEGNNLQOLVLTICQAHFAPISPGFSRELHSLISOLF 240
QY 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIQEEFSHMLICRAGAPASRHAGKVQCK 300
DB 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVIQEEFSHMLICRAGAPASRHAGKVQCK 300
QY 301 IQKVRFGCKPPRSRSVPIKRNAILHRNEWRPPAGAKARSIKMIERPKIAAVCGHYD 360
DB 301 IQKVRFGCKPPRSRSVPIKRNAILHRNEWRPPAGAKARSIKMIERPKIAAVCGHYD 360
QY 361 YVQALDMLRRRAHKSYPHPIQENTGVEDYQGETRHGSPSQWPAEYLORKEFAQYK 420
DB 361 YVQALDMLRRRAHKSYPHPIQENTGVEDYQGETRHGSPSQWPAEYLORKEFAQYK 420
QY 421 VEKOLGLRPSAEPNPNQELRSNGEPRFQELPFRKNEMKEQEWKQLEIRQOYHND 480
DB 421 VEKOLGLRPSAEPNPNQELRSNGEPRFQELPFRKNEMKEQEWKQLEIRQOYHND 480
QY 481 MKEIRKMGREPEENSISKHTYLVKSNLPVHODASEGEAPVODIEKDLKQMLQNTKE 540
DB 481 MKEIRKMGREPEENSISKHTYLVKSNLPVHODASEGEAPVODIEKDLKQMLQNTKE 540
QY 541 SKNPEQKYKAKGVKFEINLDCISDENILQEEAMDI PNETLTFFDGMKFEYECVKEH 600
DB 541 SKNPEQKYKAKGVKFEINLDCISDENILQEEAMDI PNETLTFFDGMKFEYECVKEH 600
QY 601 GDYTDKAFKHLHCPAAGFSTQ 622
DB 570 GDYTDKAFKHLHCPAAGFSTQ 591
```

RESULT 5

```
US-10-114-270-188
; Sequence 188, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
```

```
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Zhaocong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glenda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigaru, Muralichara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liette, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: NO. US20040030110A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
```

```
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 188
; LENGTH: 403
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-270-188

Query Match      58.4%; Score 2111.5; DB 12; Length 403;
Best Local Similarity 99.0%; Pred. No. 1.2e-139;
Matches 399; Conservative 2; Mismatches 1; Indels 1; Gaps 1;

QY 1 MDKYDVIKAGGAFGKAYLAGKSDSKHCVKIENFEKMPQOEKASKEVILKMKH 60
DB 1 MDKYDVIKAGGAFGKAYLAGKSDSKHCVKIENFEKMPQOEKASKEVILKMKH 60

QY 61 PNIVAFNSFQNGRLFIYMEYCDGDLMKRINRQGVLFSDQILGWVQISLGLKH 120
DB 61 PNIVAFNSFQNGRLFIYMEYCDGDLMKRINRQGVLFSDQILGWVQISLGLKH 120

QY 61 PNIVAFNSFQNGRLFIYMEYCDGDLMKRINRQGVLFSDQILGWVQISLGLKH 120
DB 61 PNIVAFNSFQNGRLFIYMEYCDGDLMKRINRQGVLFSDQILGWVQISLGLKH 120

QY 121 DRKILHRDIAKQNIIFLSKNGWAKLGDGFIARVLNNSMELARTCTGTPTYSPEICQKP 180
DB 121 DRKILHRDIAKQNIIFLSKNGWAKLGDGFIARVLNNSMELARTCTGTPTYSPEICQKP 180

QY 181 YNNKTDIWSLGCYVLELCTLKHPFEGNNLQQLVLIKQAHFAPISPGFSRELHLSQLF 240
DB 181 YNNKTDIWSLGCYVLELCTLKHPFEGNNLQQLVLIKQAHFAPISPGFSRELHLSQLF 240

QY 241 QVSPDRPSINSILKRPFLNLI PKYLTPEV-IQEFSSHMLICRAGAPASHAGKVVQKC 299
DB 241 QVSPDRPSINSILKRPFLNLI PKYLTPEVFESEFSSHMLICRAGAPASHAGKVVQKC 300

QY 300 KIQKRFQCKPSPRGRISVPIKENAILHNRWPPAGAKARSIRKMRPKIAAVCGHYD 359
DB 301 KIQKRFQCKPSPRGRISVPIKENAILHNRWPPAGAKARSIRKMRPKIAAVCGHYD 360

QY 360 YYYAQLDMLRRRAHPSYHPIQENTGVEDYQCTRHGSPSQ 402
DB 361 YYYAQLDMLRRRAHPSYHPIQENTGVEDYQCTRHGSPSQ 403

RESULT 6
US-10-114-270-186
; Sequence 186, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malvankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Zhaozhong
; APPLICANT: Gusev, Vladimir Y.

; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glenda E.
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shimkets, Richard A.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Taupier Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liette, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; APPLICANT: Liette, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: No. US20040030110A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 186
; LENGTH: 399
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-270-186

Query Match      58.3%; Score 2106.5; DB 12; Length 399;
Best Local Similarity 99.3%; Pred. No. 2.6e-139;
Matches 399; Conservative 0; Mismatches 0; Indels 3; Gaps 1;

QY 1 MDKYDVIKAGGAFGKAYLAGKSDSKHCVKIENFEKMPQOEKASKEVILKMKH 60
DB 1 MDKYDVIKAGGAFGKAYLAGKSDSKHCVKIENFEKMPQOEKASKEVILKMKH 57

QY 61 PNIVAFNSFQNGRLFIYMEYCDGDLMKRINRQGVLFSDQILGWVQISLGLKH 120
DB 58 PNIVAFNSFQNGRLFIYMEYCDGDLMKRINRQGVLFSDQILGWVQISLGLKH 117

QY 121 DRKILHRDIAKQNIIFLSKNGWAKLGDGFIARVLNNSMELARTCTGTPTYSPEICQKP 180
DB 118 DRKILHRDIAKQNIIFLSKNGWAKLGDGFIARVLNNSMELARTCTGTPTYSPEICQKP 177

QY 181 YNNKTDIWSLGCYVLELCTLKHPFEGNNLQQLVLIKQAHFAPISPGFSRELHLSQLF 240
DB 178 YNNKTDIWSLGCYVLELCTLKHPFEGNNLQQLVLIKQAHFAPISPGFSRELHLSQLF 237
```

```
QY 241 QVSPDRPSINSILKPPFLENLPKYLTPVIOEESHMLICRAGAPASHRACKVQCK 300
Db 238 QVSPDRPSINSILKPPFLENLPKYLTPVIOEESHMLICRAGAPASHRACKVQCK 297
QY 301 IQVRFQGGPPRSRISVPIKRNAILHRNEWRPPAGAKARSIMKIERKIAAVCGHYDY 360
Db 298 IQVRFQGGPPRSRISVPIKRNAILHRNEWRPPAGAKARSIMKIERKIAAVCGHYDY 357
QY 361 YYAQLDMLRRRAKPSYHPPIQENTGVEDYGQTRHGPPSQ 402
Db 358 YYAQLDMLRRRAKPSYHPPIQENTGVEDYGQTRHGPPSQ 399

RESULT 7
US-10-162-706-5
; Sequence 5, Application US/10162706
; Publication No. US20030059918A1
; GENERAL INFORMATION:
; APPLICANT: Smolyar, Alex
; TITLE OF INVENTION: Regulation of Human Serine/Threonine
; FILE OF INVENTION: Kinase
; FILE REFERENCE: 004974.00731
; CURRENT APPLICATION NUMBER: US/10/162,706
; CURRENT FILING DATE: 2002-06-06
; PRIOR APPLICATION NUMBER: US 60/296,164
; PRIOR FILING DATE: 2001-06-07
; PRIOR APPLICATION NUMBER: US 60/323,100
; PRIOR FILING DATE: 2001-09-19
; PRIOR APPLICATION NUMBER: US 60/330,578
; PRIOR FILING DATE: 2001-10-25
; PRIOR APPLICATION NUMBER: US 60/348,601
; PRIOR FILING DATE: 2002-01-17
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 5
; LENGTH: 774
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-10-162-706-5

Query Match 30.6%; Score 1106.5; DB 14; Length 774;
Best Local Similarity 35.2%; Pred. No. 7.2e-69;
Matches 276; Conservative 121; Mismatches 207; Indels 181; Gaps 23;

QY 1 MDKYDVIKAIQGAFAKAYLAKGSDSKHCIVKINFEKMPIOEKESKKEVILEKMKH 60
Db 1 MEKYVRLQKIGEGSFGKAYLVKSTEDGRHYVKEINISRMSSDKERQESREAVLANMKH 60
QY 61 PNIVAFNFSQENGRFLIVMEYCDGDLMKRINRQGVLFSEDOILLGWVQISLGLKHH 120
Db 61 PNIVQKESFEENGSLYIVMDYCEGDLFRKINAQKALFOEDQILDWVQICLAKHHV 120
QY 121 DRKILHRDIKAQNIPLSKNGMVAKLGDGFIARVYLNNSMELARTICITPYLSPICQNK 180
Db 121 DRKILHRDIKSONIFLTQDGTV-QLGDFGIARVNSTVELARTICITPYLSPICENKP 179
QY 181 YNNKTDIWSLGCVLIELCTLKHPEGNLQOLVILKICQAHFAPISPGFSREHLSLSOLF 240
Db 180 YNNKSDIWAAGCVLYELCTLKHAFEGAKNMKNLVKISGSPFVPSPHYSLRSLLSOLF 239
QY 241 QVSPDRPSINSILKPPFLENLPKYLTPVIOEESHMLICRAG----- 285
Db 240 KRNPRDRPSVNSILEKGFIAKRIEKLFLSPOLIAEEFCLTSLKFGPQLPGKRPASGQGV 299
QY 286 -----APASRIRAGKVQVCKTKQVRFQCKCPPRSRIS-VPIKRNAILHRNEWRP 333
Db 300 SSFVPAQKITPKAAKTYKGYVPLTYKYGDKLLEKPPPKHQAHQIPVK- ---MNSGEERK 356
QY 334 PAG--AQKARSIMKIER----- 348
Db 357 KMSEBAKRRLEFIKEKKQKQDIIRFLKAEQMKRQEKORLERINRAREQGNVLRAGG 416
```

```
QY 349 -----PKIAAVCGHYDYVYAQLDMLRR-RAH-----KPSYHP--IPOE 383
Db 417 SGVEKASFPFGGAVSPSPCPRQGYEHYHAIFQMORLRAEDNEARWKGGIYGRWLDPR 476
QY 384 NTG---VEDYQGETRHGFPSPQWPAEYLQKXFEAQYQKLVKE-----KQLGLRPS 431
Db 477 QKGHLAVERAQ-----VEEFQKRKEAMQNKARAEHGVVYLARLQIRLQ--- 522
QY 432 AEPNAYNROELR-----SNGEEPFRQELPRKKNEMKEQYKQLEELRQOVHN 479
Db 523 ---NFNERQQIKAKLRGENKEADGTKGQEAT-BETDNR---LKKWESLKAOTNARAAVLK 575
QY 480 DMKIRKMGREPBENSISKIYLVKSNLFPVHQDASE-GEAPVQDIEKDKQMLQNT 538
Db 576 EQLERKRKEAYERBKVWEHLVARVKSDDVPLPLELLETGSP----- 619
QY 539 KESKNPEQKYAKGVKVFENLDRKISDENLILQEEBAMDIPNETLT-----FEDGMK 590
Db 620 ---SKQVKPVISVTSALKEVGLDGLTD---TQEEEMKSNISAISKREILRLNENLK 673
QY 591 FKEYECVKEHGDYTDKAFKELHCPAEGFSTQTVAAVGNRRQWDCGAPQTLIQMMVADIT 650
Db 674 AQEDEKEKQH--HSGSCETVGHKDEREYETEN-AISSDRKKWEMGG-QLVPLDVLDT 729
QY 651 STCPT 655
Db 730 SFSAT 734

RESULT 8
US-09-783-320-4
; Sequence 4, Application US/09783320
; Patent No. US20020038011A1
; GENERAL INFORMATION:
; APPLICANT: Walke, D. Wade
; APPLICANT: Hu, Yi
; APPLICANT: Nepomnichy, Boris
; APPLICANT: Turner, C. Alexander Jr
; APPLICANT: Zambrowicz, Brian
; TITLE OF INVENTION: NO. US20020038011A1el Human Kinases and Polynucleotides Encoding
; FILE REFERENCE: LEX-0137-USA
; CURRENT APPLICATION NUMBER: US/09/783,320
; CURRENT FILING DATE: 2001-02-15
; PRIOR APPLICATION NUMBER: US 60/183,582
; PRIOR FILING DATE: 2000-02-16
; PRIOR APPLICATION NUMBER: US 60/184,014
; PRIOR FILING DATE: 2000-02-22
; NUMBER OF SEQ ID NOS: 50
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 1214
; TYPE: PRT
; ORGANISM: homo sapiens
; US-09-783-320-4

Query Match 29.9%; Score 1081.5; DB 9; Length 1214;
Best Local Similarity 34.4%; Pred. No. 7.2e-67;
Matches 274; Conservative 129; Mismatches 226; Indels 167; Gaps 21;

QY 1 MDKYDVIKAIQGAFAKAYLAKGSDSKHCIVKINFEKMPIOEKESKKEVILEKMKH 60
Db 1 MEKYVRLQKIGEGSFGKAYLVKSTEDGRHYVKEINISRMSSDKERQESREAVLANMKH 60
QY 61 PNIVAFNFSQENGRFLIVMEYCDGDLMKRINRQGVLFSEDOILLGWVQISLGLKHH 120
Db 61 PNIVQYRESFENGSLYIVMDYCEGDLFRKINAQKALFOEDQILDWVQICLAKHHV 120
QY 121 DRKILHRDIKAQNIPLSKNGMVAKLGDGFIARVYLNNSMELARTICITPYLSPICQNK 180
Db 121 DRKILHRDIKSONIFLTQDGTV-QLGDFGIARVNSTVELARTICITPYLSPICENKP 179
QY 181 YNNKTDIWSLGCVLIELCTLKHPEGNLQOLVILKICQAHFAPISPGFSREHLSLSOLF 240
```


Db 121 DRKILHDIKSONIFLTKDGTG-QLGDFGIARVLNSTVELARTCTGTGTPPYLSPEICENKP 179
QY 181 YNNKTDIWSLGLCYELCTLKHFFEGNNLQQLVLKICQAHFAPISPGFSRELHLSLSOLF 240
Db 180 YNNKSDIWLGLCYELCTLKHFAFAGNNKNVLKLIISGSPFPVSPHYSDYLSLSOLF 239
QY 241 QVSPDRPSINSILKPPFLNIPKVLTPVTCFESHMLICRAG----- 285
Db 240 KRNPRDRPSVNSILEKGFIAKIEKFLSPQLAEBFCLTKLSKFGPQLPGKRPSAGGV 299
QY 286 -----APASRHAGVWOKCKIQKVFQCKPPRSRIS--VPIKRNAILHNEWRP 333
Db 300 SSFVPAQKTKPAKYGVLTKYKYGDKLLEKPPKHKQAHQIPVK---MNSGEERK 356
QY 334 PAG-AQKARSIMIER----- 348
Db 357 KMSEEAARKRLLEFIEKEKKQDIQIRFLKAEQMKRQKQRLRINRAREQGRNVLRAGG 416
QY 349 -----PKIAAVCGHYDYVAQDLMLRR-RAH-----KPSVHP-IPQE 383
Db 417 SGEVKASFFGIGGAVSPSCSPRQVQYEHYHALFDQMORLAEDNEARWKGGIYGRWLPER 476
QY 384 NTG---VEDYGBQTRHGPSQWPAFYLQKPEAQYKLVK-----KOLGLRPSS 431
Db 477 QXGHLAVERANQ-----VEEFLQKREAMQNKARAGHVYVLARLQIRLQ--- 522
QY 432 AEPYNNORQELR-----SNGEPRQELPPFRKNEMKEQYKQLEBEIFQOYHN 479
Db 523 ---NFNERQOIKAKLGENKEADGTGQEAT-EETDMR---LKKMESLKAQTNARAVALK 575
QY 480 DMKEIRKMGREPEENSKISHKTYLVKKNLPV 512
Db 576 EQLERKRKEAYEREKKVWEHLVARVKSSDVPL 608

RESULT 11

US-09-898-837A-35
; Sequence 35, Application US/09898837A
; Publication No. US2003007697A1
; GENERAL INFORMATION:
; APPLICANT: Quinn, Kerry E.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Majumder, Kumud
; APPLICANT: Vernet, Corine
; APPLICANT: Herrmann, John L.
; APPLICANT: Burgess, Catherine
; APPLICANT: Fernandes, Elma
; APPLICANT: Taupier, Jr., Raymond
; APPLICANT: Rastelli, Luca
; APPLICANT: Curagen Corporation
; APPLICANT: Gerlach, Valerie L.
; APPLICANT: MacDougall, John R.
; TITLE OF INVENTION: NOVEL SERINE/THREONINE PROTEIN-KINASE LIKE PROTEINS AND
; FILE REFERENCE: 15966-598 CIP
; CURRENT APPLICATION NUMBER: US/09/898, 837A
; CURRENT FILING DATE: 2001-07-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/165,986
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,839
; PRIOR FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/195,637
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/197,080
; PRIOR FILING DATE: 2000-04-13
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/232,677
; PRIOR FILING DATE: 2000-09-15
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/181,347
; PRIOR FILING DATE: 2000-02-09
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/194,195
; PRIOR FILING DATE: 2000-04-03
; PRIOR APPLICATION NUMBER: U.S.S.N. 60/215,906
; PRIOR FILING DATE: 2000-07-03

; PRIOR APPLICATION NUMBER: U.S.S.N. 09/715,427
; PRIOR FILING DATE: 2000-11-16
; NUMBER OF SEQ ID NOS: 53
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 35
; LENGTH: 254
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-898-837A-35
Query Match 25.6%; Score 926.5; DB 10; Length 254;
Best Local Similarity 66.8%; Pred. No. 7.2e-57;
Matches 169; Conservative 41; Mismatches 42; Indels 1; Gaps 1;
QY 7 IKATGQAGKAYLAKSKSDSKVCIKENINFERMPTIOEASKEVILLKMKHPNIVAF 66
Db 3 LQKTGEGSFGKAVLVKSTEDGRHYVKEINISMSDKERQESRREVAVLANMKHPNIVQY 62
QY 67 FNSFOENGRIFIVNEYCDGDLMKRINRQRGVLFSEDOILGWVQVLSGLKHIHDKILH 126
Db 63 KESFEENGSLYVNDYCEGGDLFKRINAQKALFQEDQILDWFVQICLALKVHDKILH 122
QY 127 RDIKAQNFUSKGNWAKLGDGFIARVLNNSMELARTCTGTPTPYLSPEICQNPYNNKTD 186
Db 123 RDIKSONIFLTKDGTG-QLGDFGIARVLNSTVELARTCTGTPTPYLSPEICENKFNKSD 181
QY 187 IWSLGCVLVELCTLKHFFEGNNLQQLVLKICQAHFAPISPGFSRELHLSLSOLFQVSPDR 246
Db 182 IWLGCVLVELCTLKHFAFAGNNKNVLKLIISGSPFPVSPHYSDYLSLSOLFQVSPDR 241
QY 247 RPSINSILKRPFL 259
Db 242 RPSVNSILEKGF 254

RESULT 12

US-10-114-270-180
; Sequence 180, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
; APPLICANT: Miller, Charles E.
; APPLICANT: Malyankar, Uriel M.
; APPLICANT: Spytek, Kimberly A.
; APPLICANT: Patturajan, Meera
; APPLICANT: Liu, Ziaohong
; APPLICANT: Gusev, Vladimir Y.
; APPLICANT: Li, Li
; APPLICANT: Vernet, Corine
; APPLICANT: Zerhusen, Bryan D.
; APPLICANT: Gorman, Linda
; APPLICANT: Shenoy, Suresh G.
; APPLICANT: Pena, Carol E.A.
; APPLICANT: Smithson, Glennnda
; APPLICANT: Burgess, Catherine E.
; APPLICANT: Gerlach, Valerie
; APPLICANT: Padigaru, Muralidhara
; APPLICANT: Shinkets, Richard A.
; APPLICANT: Gangolli, Esha A.
; APPLICANT: Taupier, Jr., Raymond J.
; APPLICANT: Casman, Stacie J.
; APPLICANT: Ji, Weizhen
; APPLICANT: Anderson, David W.
; APPLICANT: Liote, Mario W.
; APPLICANT: Rastelli, Luca
; APPLICANT: Edinger, Shlomit R.
; APPLICANT: Stone, David J.
; APPLICANT: MacDougall, John R.
; APPLICANT: Rothenberg, Mark E.
; TITLE OF INVENTION: NO. US20040030110A1el Proteins and Nucleic Acids Encoding Same
; FILE REFERENCE: 21402-322C
; CURRENT APPLICATION NUMBER: US/10/114,270

```
; CURRENT FILING DATE: 2002-11-27
; PRIOR APPLICATION NUMBER: 60/281,086
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,136
; PRIOR FILING DATE: 2001-04-03
; PRIOR APPLICATION NUMBER: 60/281,863
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/281,906
; PRIOR FILING DATE: 2001-04-05
; PRIOR APPLICATION NUMBER: 60/282,020
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: 60/282,930
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/282,934
; PRIOR FILING DATE: 2001-04-10
; PRIOR APPLICATION NUMBER: 60/283,512
; PRIOR FILING DATE: 2001-04-12
; PRIOR APPLICATION NUMBER: 60/283,710
; PRIOR FILING DATE: 2001-04-13
; PRIOR APPLICATION NUMBER: 60/284,234
; PRIOR FILING DATE: 2001-04-17
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 470
; SEQ ID NO 180
; LENGTH: 506
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-114-270-180

Query Match      22.6%; Score 818; DB 12; Length 506;
Best Local Similarity 34.4%; Pred. No. 6.9e-49;
Matches 198; Conservative 97; Mismatches 178; Indels 102; Gaps 16;

QY 1 MDKYDVKAIGQAFKAYLAKGSDSKHCIVKEINFEMPTQOEKASKEVILEKMKH 60
Db 1 MDDYVLMRIGSGFGRALLVQHSSNQWFMKEIRLPKVTNTQN-SRKEAVLLAKMKH 59

QY 61 PNIVAFNSFOENGRFLFVMEYCDGDLMKRNROGVLPFSDQILGWVQISLGKHH 120
Db 60 PNIVAFKSEAEGLHYIYMEYCDGDLMKQIKQKQKGLFPEDQILNFTQCLGVNHH 119

QY 121 DRKILHRDIKAQNIPLSKNGWAKLDFGFIARVLNNSMELARTCTGTPTVYLSPEICNKP 180
Db 120 KKEVLHRDIKSQNIPLTQNGKV-KLGDGFSARLLSNPMAFACTYGTPTVYLPPEINLP 178

QY 181 YNNKTDIWSLGCVLVELCTLKHPFEGNNLQQLVLKICQAHFAPISPGFSRELHLSIQLP 240
Db 179 YNNKSDIWSLGCILVELCTLKHPFQANGWKNLILKVCQCCISPLSPSHYSYELQFLVKQMF 238

QY 241 QVSPDRPDSINSILKPPFLNLIIPKYLTPETVIOEFSHMLICRAGAPASRHAGKVQCK 300
Db 239 KRNPSHRPSATILSGRIVARLVQKCLPEIINE-YGEEVIEIK 282

QY 301 IQKVRPQGCPPRSRISVPIKRNAILHRNEWPPAGAAQKARSIKMIERPRTAAVCGHYDY 360
Db 283 NKGHTNPKTKTPNRIRIALNGEASTVQEEQDRKGS----- 319

QY 361 YYAQDMLRRRAHKSYPHPPOENTGVEDYQOETRHGSPS-----QWPAEYLQRKFEAQ 416
Db 320 -HTDESINENIUVESALRRVNEEG--NKSVHLRKASSPNLHRQW----- 363

QY 417 YKLKVEKQIQLRPSAEPPNYNQBELRSN--GEEPR-FQELPPFRKNEMKEQYKQLEEI 473
Db 364 -----EKNV-----PNTALTALENASILTSSLTAEDDRGGGVIKYSKNTTRKQ--W-LKET 411

QY 474 RQYVNDMKEIRKNGRPEENSLSKSHKTYLV-----KKSNIPLVHQDASEG-----E 520
Db 412 PTLNLNLK-----NADLSAFQTYTYRPGSEGFLKGPLSEETASDSVDGHD 461

QY 521 APVQD---IEKDLKQWLQNTKESNPKQYKAKK 552
Db 462 SVILDPERLEPLGDEEDTDDEEDNDPDWVSELKK 496
```

```
RESULT 13
US-09-870-962-6
; Sequence 6, Application US/09870962
; Patent No. US20020081290A1
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Valda
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/870,962
; CURRENT FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: 09/420,915
; PRIOR FILING DATE: 1999-10-20
; PRIOR APPLICATION NUMBER: US 09/173,581
; PRIOR FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 1567782
US-09-870-962-6

Query Match      22.6%; Score 816.5; DB 9; Length 345;
Best Local Similarity 44.9%; Pred. No. 5.4e-49;
Matches 173; Conservative 55; Mismatches 80; Indels 77; Gaps 6;

QY 1 MDKYDVKAIGQAFKAYLAKGSDSKHCIVKEINFEMPTQOEKASKEVILEKMKH 60
Db 1 MEKYVRLQKIGEGSGKAILVKSTEDGRQYVKEINISMSKSEESREVRVAVLANMKH 60

QY 61 PNIVAFNSFOENGRFLFVMEYCDGDLMKRNROGVLPFSDQILGWVQISLGKHH 120
Db 61 PNIVQRESF-----EGILDWVQICLALKVH 88

QY 121 DRKILHRDIKAQNIPLSKNGWAKLDFGFIARVLNNSMELARTCTGTPTVYLSPEICNKP 180
Db 89 DRKILHRDIKSQNIPLTQNGTV-QLGDGFIARVLNSTVELARTCTGTPTVYLSPEICENKP 147

QY 181 YNNKTDIWSLGCVLVELCTLKHPFEGNNLQQLVLKICQAHFAPISPGFSRELHLSIQLP 240
Db 148 YNNKSDIWSLGCVLVELCTLKHPFEGNNLQQLVLKICQAHFAPISPGFSRELHLSIQLP 207

QY 241 QVSPDRPDSINSILKPPFLNLIIPKYLTPETVIOEFSHMLICRAGAPASRHAGKVQCK 293
Db 208 KRNPRDRPDSINSILEKGFIAKRIEKLSPQIAEBECLTKTFSKFSQIPAKRPSAQNS 267

QY 294 -----KVQCKIKQVRFQGCPPRSRISVPIKRNAIL 325
Db 268 ISVMPAQKTKPAKYGIPLAYKYGDKLHKKPLQGHK-QAHQTPKRVNTGEERRKI 326

QY 326 LHRNEWPPAGAAQKARSIKMIERPRTAAVCGHYDY 350
Db 327 SEB-----AARKERLEFIEKDK 343

RESULT 14
US-10-114-270-184
; Sequence 184, Application US/10114270
; Publication No. US20040030110A1
; GENERAL INFORMATION:
; APPLICANT: Guo, Xiaojia
; APPLICANT: Kekuda, Ramesh
```



```

; SEQ ID NO 182
; LENGTH: 489
; TYPE: FRT
; ORGANISM: Homo sapiens
US-10-114-270-182

Query Match      22.5%; Score 812; DB 12; Length 489;
Best Local Similarity 54.7%; Pred. No. 1.7e-48;
Matches 150; Conservative 52; Mismatches 70; Indels 2; Gaps 2;

Qy 1 MDKYDVKAIGQAGFKAYLAKGSDSKHCVIKEINFERMPIQEKASKEVILLKMKH 60
Db 1 MDDYVLRMIGEGSGRALLVQHESNQMPAMKEIRLPKVTNTQN-SRKEAVLLAKMKH 59

Qy 61 PNIVAFNFSFOENGRIFIVMEYCDGDLMKRINRQGVLFSEDOILGHFVQISGLKHIH 120
Db 60 PNIVAFNFSFOENGRIFIVMEYCDGDLMKRINRQGVLFSEDOILGHFVQISGLKHIH 119

Qy 121 DRKILHRDIKAQNIFLSKNGMYAKLGDGFIARVLNNSMELARTCICTPYILSPEICONKP 180
Db 120 KKEVLHRDIKSNIFLTQNGV-KLGDGFSARLLSNPMAFACTYVGTPTYVPEIWEINLP 178

Qy 181 YNNKTDINSLGCVLYELCTLKHPFEGNNLQQLVLCQAHFAPISFGFSRELHLSLSOLF 240
Db 179 YNNKTDINSLGCVLYELCTLKHPFEGNNLQQLVLCQAHFAPISFGFSRELHLSLSOLF 238

Qy 241 QVSPDRPSINSILKRPFLLENIPKYLTPETVIOE 274
Db 239 KNPFSRPSATILLSRGIVARLVQKCLPPEIIME 272

```

Search completed: October 8, 2004, 17:20:45
Job time : 139 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2004, 17:00:40 ; Search time 62 Seconds
(without alignments)
3112.580 Million cell updates/sec

Title: US-09-940-921B-2
Perfect score: 3616
Sequence: 1 MDKYDVIRKAGGAFKAYL.....SVSRQEGTKDPYSPVLILM 683

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_29Jan04:*
1: Genesecp1980s:*
2: Genesecp1990s:*
3: Genesecp2000s:*
4: Genesecp2001s:*
5: Genesecp2002s:*
6: Genesecp2003as:*
7: Genesecp2003bs:*
8: Genesecp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	3616	100.0	683	5 AAU77928	Amino aci
2	3431.5	94.9	654	5 AAU77929	Amino aci
3	3081.5	85.2	640	6 ABG72000	Human ser
4	2925	80.9	649	4 AAU03545	Human ser
5	2820	78.0	631	6 ABP71710	Human NEK
6	2111.5	58.4	403	6 ABU54635	Human NOV
7	2106.5	58.3	399	6 ABU54634	Human NOV
8	1118.5	30.9	544	6 ABP96073	Human pro
9	1106.5	30.6	774	6 AAO26613	Serine/th
10	1106.5	30.6	774	6 ABP71711	NEK-like
11	1106.5	30.6	774	6 AAO16441	Human ser
12	1106.5	30.6	774	6 ABP97691	Amino aci
13	1105.5	30.6	1242	4 ABP39210	Human pol
14	1105.5	30.6	1242	6 AAO16440	Human ser
15	1081.5	29.9	1214	4 AAU07102	Human nov
16	1077.5	29.8	1214	4 ABP39211	Human pol
17	1077.5	29.8	1214	6 ABU08113	Human kin
18	930.5	25.7	632	5 ABP55160	Ser/Thr/T
19	818	22.6	506	6 ABU54631	Human NOV
20	816.5	22.6	345	3 ABU76753	Human pro
21	816.5	22.6	345	4 AAO08211	Human pro
22	816.5	22.6	345	5 ABP84438	Human pro
23	812.5	22.5	546	6 ABP96072	Human pro
24	812	22.5	425	6 ABU54633	Human NOV
25	812	22.5	489	6 ABU54632	Human NOV

26	808.5	22.4	511	6 ABP71712	NEK-like
27	808.5	22.4	511	6 ABP97690	Amino aci
28	808	22.3	506	4 AAM78344	Human pro
29	808	22.3	506	5 ABP60668	Human ser
30	806	22.3	506	5 AAE24136	Human kin
31	804.5	22.2	507	6 ABP97688	Amino aci
32	804	22.2	527	4 AAM79328	Human pro
33	804	22.2	527	5 ABP97328	Human pro
34	800	22.1	489	7 ABB97224	Novel hum
35	741.5	20.5	510	3 AAY68778	Human KPP
36	728.5	20.1	459	6 ABP71713	Amino aci
37	728.5	20.1	459	6 ABP97689	NEK-like
38	668	18.5	841	7 ADB37570	Neural th
39	660	18.3	841	7 ADB37605	Human ser
40	636.5	17.5	329	6 AAO26614	Human ser
41	634.5	17.5	713	6 AAO26615	Human ser
42	634	17.5	692	5 AAO15418	Human ser
43	632.5	17.5	280	6 AAO26612	Novel hum
44	632.5	17.5	460	5 AAE21728	Human ser
45	632.5	17.5	460	6 AAO26616	Human PKI
					Amino ser

ALIGNMENTS

RESULT 1
AAU77928
ID AAU77928 standard; protein; 683 AA.
XX
AC AAU77928;
XX
DT 18-JUN-2002 (first entry)
XX
DE Amino acid sequence for novel human kinase protein #1.
XX
KW Novel human protein; NHP; serine-threonine kinase; brain;
KW calcium/calmodulin-dependent kinase; myosin light chain kinase;
KW biological disorder; spleen; placenta; chromosome 6; enzyme.
XX
OS Homo sapiens.
XX
FN WO200218555-A2.
XX
PD 07-MAR-2002.
XX
PF 28-AUG-2001; 2001WO-US026776.
XX
PR 31-AUG-2000; 2000US-0229280P.
XX
PI (LEXI-) LEXICON GENETICS INC.
XX
PI Fiddle CJ, Hilbun E, Nepomnichy B, Hu Y;
DR WPI; 2002-292200/33.
DR N-PSDB; ABK12424.
XX
XX Novel polynucleotide encoding novel human protein sharing structural similarity with animal kinases e.g. serine-threonine, calcium/calmodulin-dependent, and myosin light chain kinases, useful as probes and primers.
XX
XX Claim 1; Page 38-39; 46pp; English.
XX
XX The present invention relates to the isolation of novel human proteins (NHPs) and the polynucleotide sequences encoding them. The NHPs of the invention are kinase proteins and share structural similarity to serine-threonine, calcium/calmodulin-dependent, and myosin light chain kinases. The sequences of the invention are useful for treating biological disorders. The polynucleotide sequences encoding the kinase proteins can be used as primers and probes. The sequences are also useful for identifying mutations associated with a particular disease and also in a prognostic or diagnostic assay. The present sequence represents human protein kinase #1 which is expressed in a broad range of human tissues such as brain, spleen, and placenta. The gene encoding protein kinase #1

Db 465 -----ENSKISHKTLVKKSNLPVHQDASEGAPVQIEKDLQKRLQNTKE 511
QY 541 SKNPEQKYKAKGVKFEINLDCISDENILQEEAWDIPNETLTPEDEGMKFEYECVKEH 600
Db 512 SKNPEQKYKAKGVKFEINLDCISDENILQEEAWDIPNETLTPEDEGMKFEYECVKEH 571
QY 601 GYTDKAFKELHCPAGFSTQTVAAVGNRRQDGGAPQTLLOWMAVADITSTCTGPDSE 660
Db 572 GYTDKAFKELHCPAGFSTQTVAAVGNRRQDGGAPQTLLOWMAVADITSTCTGPDSE 631
QY 661 SVLSVSRQEGKTKDPYSPVLILM 683
Db 632 SVLSVSRQEGKTKDPYSPVLILM 654
RESULT 3
ABG72000
ID ABG72000 standard; protein; 640 AA.
XX
AC ABG72000;
XX
DT 04-FEB-2003 (first entry)
XX
DE Human serine/threonine kinase.
XX
KW Human; enzyme; serine/threonine kinase; lung carcinoma; chromosome 13.
XX
OS Homo sapiens.
XX
PN WO200281727-A2.
XX
PD 17-OCT-2002.
XX
PF 02-APR-2002; 2002WO-US010156.
XX
PR 03-APR-2001; 2001US-00824583.
XX
PR 05-JUN-2001; 2001US-00873404.
XX
PA (PEKE) PE CORP NY.
XX
PI Webster M, Yan C, Di Francesco V, Beasley EM;
XX
DR WPI; 2003-058562/05.
XX
DR N-PSDB; ABS57149, ABS57150.
XX
PT Novel human kinase protein expressed in lung carcinoma and placenta is
PT useful to diagnose and treat diseases and disorders associated with
PT expression or activity of the protein.
XX
PS Claim 1; Fig 2A; 101pp; English.
XX
CC The invention relates to an isolated human kinase peptide of the
CC serine/threonine, an allelic variant or orthologue, and encoded by a
CC nucleic acid that hybridises under stringent conditions with the cDNA and
CC gene sequences appearing as ABS57149 and ABS57150, or a fragment
CC comprising at least 10 contiguous amino acids. Also included are an
CC antibody that selectively binds to the kinase, a gene chip comprising the
CC nucleic acids, a transgenic non human animal comprising the nucleic
CC acids, a nucleic acid vector comprising the nucleic acids, a host cell
CC containing the vector and expressing the kinase, identifying a
CC modulator/binding agent of the kinase (comprising contacting the peptide
CC with an agent and determining if the agent has modulated function,
CC expression or activity of the peptide or formed a complex with it), The
CC molecules of the invention are useful to diagnose and treat a disorder
CC characterised by aberrant expression of the protein (e.g. lung
CC carcinoma). Agents which modulate the function or activity of the protein
CC are useful to treat diseases or disorders mediated by human kinase
CC protein. The gene for encoding the novel kinase is located on chromosome
CC 13. The present sequence represents the novel human kinase
XX
SQ Sequence 640 AA;

Query Match 85.2%; Score 3081.5; DB 6; Length 640;
Best Local Similarity 94.1%; Pred. No. 8.1e-233;
Matches 585; Conservative 2; Mismatches 4; Indels 31; Gaps 1;
QY 1 MDKVDVTKAQGAGFAGKAYLAKGSDSKHCVIIEINPEKMPIOEKEASKKEVILLEKWKH 60
Db 1 MDKVDVTKAQGAGFAGKAYLAKGSDSKHCVIIEINPEKMPIOEKEASKKEVILLEKWKH 60
QY 61 PNIVAFNSFOENGRLFIIVMEYCDGDLMKRINRQGVLFSEDIQLGFWFQISLGLKHH 120
Db 61 PNIVAFNSFOENGRLFIIVMEYCDGDLMKRINRQGVLFSEDIQLGFWFQISLGLKHH 120
QY 121 DRKILHRDIKAQNTFLSKNGWAKLDFGIARVLNNSMELARTICIGTPYLSPEICQKP 180
Db 121 DRKILHRDIKAQNTFLSKNGWAKLDFGIARVLNNSMELARTICIGTPYLSPEICQKP 180
QY 181 YNNKTDIWSLGCVLVELCTLKHPPENNLQOLVLIQCAHFAPISPGFSRSLHSLSOLF 240
Db 181 YNNKTDIWSLGCVLVELCTLKHPPENNLQOLVLIQCAHFAPISPGFSRSLHSLSOLF 240
QY 241 QVSPDRPSINSILKRPFLNLPKYLTPTEVIOEFESHMLICRAGAPASHAGKVKQCK 300
Db 241 QVSPDRPSINSILKRPFLNLPKYLTPTEVIOEFESHMLICRAGAPASHAGKVKQCK 300
QY 301 IQKVRFOGKCPPRSISVPIKSNAILHRNWRPPAGAKARSIKMIERPKIAAVCGHYD 360
Db 301 IQKVRFOGKCPPRSISVPIKSNAILHRNWRPPAGAKARSIKMIERPKIAAVCGHYD 360
QY 361 YYAQLDMLRRRAHKPSYHPPIQENTGVEDYQGETRHGSPSPQWPAEYLQKFEAQYK 420
Db 361 YYAQLDMLRRRAHKPSYHPPIQENTGVEDYQGETRHGSPSPQWPAEYLQKFEAQYK 420
QY 421 VEKQLGLPSSAEPNPNQRLNSGEBPRPQELPFRKNEMKEQYWKQLEIEIQOYHND 480
Db 421 VEKQLGLPSSAEPNPNQRLNSGEBPRPQELPFRKNEMKEQYWKQLEIEIQOYHND 480
QY 481 MKEIRKMGREPE-----DIEKDLQKQMLQNTKE 540
Db 481 MKEIRKMGREPE-----DIEKDLQKQMLQNTKE 509
QY 541 SKNPEQKYKAKGVKFEINLDCISDENILQEEAWDIPNETLTPEDEGMKFEYECVKEH 600
Db 510 SKNPEQKYKAKGVKFEINLDCISDENILQEEAWDIPNETLTPEDEGMKFEYECVKEH 569
QY 601 GYTDKAFKELHCPAGFSTQTVAAVGNRRQDGGAPQTLLOWMAVADITSTCTGPDSE 622
Db 570 GYTDKAFKELHCPAGFSTQTVAAVGNRRQDGGAPQTLLOWMAVADITSTCTGPDSE 591
RESULT 4
AAU03545
ID AAU03545 standard; protein; 649 AA.
XX
AC AAU03545;
XX
DT 12-SEP-2001 (first entry)
XX
DE Human protein kinase #45.
XX
KW Human; protein kinase; PTK; STK; cancer; cardiovascular disease;
KW metabolic disorder; immune related disease; neurological disorder;
KW neurodegenerative disorder; inflammatory disorder; infectious disease;
KW reproductive disorder.
XX
OS Homo sapiens.
XX
FN WO200139503-A2.
XX
PD 31-MAY-2001.
XX
PF 22-NOV-2000; 2000WO-US032085.
XX
PR 24-NOV-1999; 99US-0167482P.

XX PA (SUGEN-) SUGEN INC.

XX PI Plowman GD, Whyte D, Manning G, Sudarsanam S, Martinez R;

XX PI Flanagan P, Clary D;

XX DR WPI; 2001-343950/36.

XX DR N-PSDB; AAS06745.

XX PT Nucleic acids encoding human kinase polypeptides, useful for preventing

XX PT diagnosing and/or treating e.g. cancer, immune, cardiovascular and

XX PT neuronal-associated diseases, and microbial infections.

XX PS Claim 7; Fig 2; 433pp; English.

XX CC AAU03501-AAU03557 represent novel human protein kinases #1-57. The novel

XX CC protein kinases have been identified as members of the tyrosine or

XX CC serine/threonine kinase (STK and STR) families. The polynucleotides

XX CC encoding protein kinases and the polypeptides may be used in the

XX CC prevention, diagnosis and treatment of diseases associated with

XX CC inappropriate kinase expression. For example, they may be used to treat

XX CC cancers (especially cancers of haematopoietic origin), cardiovascular

XX CC disease (e.g. atherosclerosis), metabolic disorders (e.g. diabetes),

XX CC immune related diseases (e.g. rheumatoid arthritis), neurological

XX CC disorders (e.g. schizophrenia), neurodegenerative disorders (e.g.

XX CC Parkinson's disease), inflammatory disorders (e.g. asthma), infectious

XX CC disease (e.g. HIV) and reproductive disorders (e.g. infertility).

XX CC Additionally, polynucleotides encoding protein kinases may be used for

XX CC gene therapy and as DNA probes in diagnostic assays. The protein kinase

XX CC polypeptides may be used as antigens in the production of antibodies

XX CC against the protein kinases and in assays to identify modulators of

XX CC protein kinase expression and activity

XX SQ Sequence 649 AA;

Query Match 80.9%; Score 2925; DB 4; Length 649;

Best Local Similarity 84.2%; Pred. No. 1.6e-220;

Matches 570; Conservative 12; Mismatches 31; Indels 64; Gaps 6;

QY 1 MDKYDVKAIGQAGAFKAYLAKGKSDSKHCVCIKENFERMPIQEKAEKKEVILLEMKH 60

DB 1 MDKYDVKAIGQAGAFKAYLAKGKSDSKHCVCIKENFERMPIQEKAEKKEVILLEMKH 60

QY 61 PNIVAFNSFOENGRFPIVMEYCDGDLNKRINRQGVLFSDQILGFWFVQISLGLKHIH 120

DB 61 PNIVAFNSFOENGRFPIVMEYCDGDLNKRINRQGVLFSDQILGFWFVQISLGLKHIH 120

QY 121 DRKILHRDIKAQNFILSKNGWAKLGFAGIARVLNNSMELARTCTGTYYLSPELCONKP 180

DB 121 DRKILHRDIKAQNFILSKNGWAKLGFAGIARVLNNSMELARTCTGTYYLSPELCONKP 180

QY 181 YNKTDIWSLGLCVLVELCTLKHPPFGNNLQQLVLIKQAHFAPISPGFSRELHLSLSQLF 240

DB 181 YNKTDIWSLGLCVLVELCTLKHPPFGNNLQQLVLIKQAHFAPISPGFSRELHLSLSQLF 240

QY 241 QVSPDRPSINSLTKRPFLENLIPKYLTPETVQEE- ---FSHMLICRAGAPASRHAGVW 296

DB 241 QVSPDRPSINSLTKRPFLENLIPKYLTPETVQEE- ---FSHMLICRAGAPASRHAGVW 296

QY 297 QKCKIQKVFQKCPSPRSISVPIKRNAILHNRWPPAGAKAFSIXMIERPKIAAVCG 356

DB 297 QKCKIQKVFQKCPSPRSISVPIKRNAILHNRWPPAGAKAFSIXMIERPKIAAVCG 356

QY 357 HYDYTYAQLDMRLRAHAKPSYHPIQENTGVEDYQETHGSPSQWPAEYLQKFEAAQ 416

DB 357 HYDYTYAQLDMRLRAHAKPSYHPIQENTGVEDYQETHGSPSQWPAEYLQKFEAAQ 416

QY 391 HYDYTYAQLDMRLRAHAKPSYHPIQENTGVEDYQETHGSPSQWPAEYLQKFEAAQ 391

DB 391 HYDYTYAQLDMRLRAHAKPSYHPIQENTGVEDYQETHGSPSQWPAEYLQKFEAAQ 391

QY 417 YLKVKEQQLGLRPSAEPNQNORQLRNGEPPFQELPPFRKNEMKEQYWKQLEERQ 476

DB 417 YLKVKEQQLGLRPSAEPNQNORQLRNGEPPFQELPPFRKNEMKEQYWKQLEERQ 476

QY 476 YLKVKEQQLGLRPSAEPNQNORQLRNGEPPFQELPPFRKNEMKEQYWKQLEERQ 476

QY 477 YHNDMKEIRKKGREPE-----DIEKDKQWRLQ 480

DB 477 YHNDMKEIRKKGREPE-----DIEKDKQWRLQ 480

DB 452 YHNDMKEIRKKGREPE-----DIEKDKQWRLQ 480

QY 537 NTKSKNPEQKYKAKGVKFEINLKDCKISDENILQEEBAMDIPNETLTTFDGMKFEYEC 596

DB 481 NTKSKNPEQKYKAKGVKFEINLKDCKISDENILQEEBAMDIPNETLTTFDGMKFEYEC 540

QY 597 VKHGDYTDKAFKELHCFEAGFSTQTTVAANGNRQWDGAPQTLLOMMAVADITSCTPTG 656

DB 541 VKHGDYTDKAFKELHCFEAGFSTQTTVAANGNRQWDGAPQTLLOMMAVADITSCTPTG 600

QY 657 PDSVSVLSVRQEGKTK 673

DB 601 PDNGQVIVIEGIPGNRK 617

RESULT 5

ABP71710

ID ABP71710 standard; protein; 631 AA.

XX AC ABP71710;

XX DT 17-APR-2003 (first entry)

XX DE Human NEK-like serine/threonine protein kinase # SEQ ID 2.

XX KW Human; NEK-like serine/threonine protein kinase; cytostatic; cardiant;

XX KW antiinflammatory; nootropic; neuroprotective; cancer; colon cancer;

XX KW cardiovascular disorder; diabetes; COPD; CNS disorder.

XX OS Homo sapiens.

XX FN WO2003000903-A2.

XX PD 03-JAN-2003.

XX PF 24-JUN-2002; 2002WO-EP006948.

XX PR 25-JUN-2001; 2001US-0300068P.

XX PR 07-DEC-2001; 2001US-0336704P.

XX PA (PARB) BAYER AG.

XX PI Xiao Y;

XX DR WPI; 2003-184051/19.

XX DR N-PSDB; ABZ59716.

XX PT New polynucleotide encoding a NEK-like serine/threonine kinase

XX PT polypeptide useful for treating diseases associated with kinase

XX PT dysfunction, e.g. cardiovascular disorders, cancer such as colon cancer,

XX PT diabetes and CNS disorders.

XX PS Claim 1 a ii; Fig 2; 149pp; English.

XX CC The invention relates to a newly isolated polynucleotide encoding an NEK-

XX CC like serine/threonine protein kinase. The activity of the polynucleotide

XX CC and polypeptide of the invention may be described as cytostatic,

XX CC cardiant, antiinflammatory, nootropic and neuroprotective. The expression

XX CC vector and reagent of the invention are useful for the preparation of a

XX CC medicament for modulating the activity of an NEK-like serine/threonine

XX CC kinase in a disease, such as cancer (e.g. colon cancer), cardiovascular

XX CC disorder, diabetes, COPD or CNS disorder. The polypeptides may also be

XX CC used to identify compounds which may act as activators or inhibitors at

XX CC the enzyme's active site, to raise specific antibodies which can block

XX CC the enzyme and effectively reduce its activity, as a bait protein in a

XX CC two-hybrid or three-hybrid assay to identify other proteins which bind to

XX CC or interact with the human NEK-like serine/threonine kinase polypeptide

XX CC and modulate its activity, and for the immunisation of mammals. The

XX CC current sequence represents the human NEK-like serine/threonine protein

XX CC kinase of the invention

XX SQ Sequence 631 AA;

Query Match		78.0%; Score 2820; DB 6; Length 631;
Best Local Similarity		79.8%; Pred. No. 2.6e-212;
Matches 552; Conservative		9; Mismatches 19; Indels 112; Gaps 5;
QY	1	MDKYDVIKAIGQAGFKAYLAKGSDSKHCVIKIEINFEKMPIQEKXASKKEVILLKMKH 60
Db	1	MDKYDVIKAIGQAGFKAYLAKGSDSKHCVIKIEINFEKMPIQEKXASKKEVILLKMKH 60
QY	61	PNIVAFNSFOENGLFIVMEYCDGDLMKRINRQGVLFSEDOILGHFWQISLGLKHIIH 120
Db	61	PNIVAFNSFOENGLFIVMEYCDGDLMKRINRQGVLFSEDOILGHFWQISLGLKHIIH 120
QY	121	DRKILHDIKAQNIIFLSKNGMAKLGDFGIARVLNNSMELARTCIGTPYILSPICQNK 180
Db	121	DRKILHDIKAQNIIFLSKNGMAKLGDFGIARVLNNSMELARTCIGTPYILSPICQNK 180
QY	181	YNNKTDIWSLGVLYELCTLKHPFFGNLQQLVLKICQAHFAPISPGFSRSLHSLSOLF 240
Db	181	YNNKTDIWSLGVLYELCTLKHPFFGNLQQLVLKICQAHFAPISPGFSRSLHSLSOLF 240
QY	241	QVSPDRPSINSILKRPFLNLI PKYLTPEVIOEBSHMLICRAGAPASRHAGKVQK - 298
Db	241	QVSPDRPSINSILKRPFLNLI PKYLTPEVIOEBSHMLICRAGAPASRHAGKVQVQHT 300
QY	299	-----CKIQKVRFGKCPKPRSRISVPIKNAILHRNEWPPAGAKAR 341
Db	301	GVRSGLSRPWAALLRLCELQPLRL-----LSQAAVECLR-----LF 337
QY	342	STKMTERPKIAVCGHYDYVYLAQDLRRRAHKPSYHPIPOENTGVEDYGOETRHGPS 401
Db	338	QTKMTERPKIAVCGHYDYVYLAQDLRRRAHKPSYHPIPOENTGVEDYGOETRHGPS 397
QY	402	QWPAEYLQKFAQOYKLVKVEKQLGRFPSSAEPNTNQRQLRSNGEERFOELPFRKNEM 461
Db	398	QWPAEYLQKFAQOYKLVKVEKQL-----421
QY	462	KEQEVYKQLEIRIQOYHNDKMEIRKNGREPEPENSISKHTYLVKSNLPVHQDASEGEA 521
Db	422	---EYWKQLEIRIQOYHNDKMEIRKNGREPE-----450
QY	522	PVQDIEKDLKQMLQNTKESKNPEQYKAKGVKFEINLDCISDENILOEEAAMDIPNE 581
Db	451	---DIEKDLKQMLQNTKESKNPEQYKAKGVKFEINLDCISDENILOEEAAMDIPNE 507
QY	582	TLTFSDGKFKPEYCKHGDYTDKAFKLCPEAGFSTQTVAAVGNRRQWDGAPQTL 641
Db	508	TLTFSDGKFKPEYCKHGDYTDKAFKLCPEAGFSTQTVAAVGNRRQWDGAPQTL 567
QY	642	QWMAVADITSTCTGPDSESVLSVSGQKTK 673
Db	568	QWMAVADITSTCTGPDNQVIVIEGIPGNRK 599
RESULT 6		
ABU54635		
ID	ABU54635	standard; protein; 403 AA.
XX	AC	ABU54635;
XX	DT	03-JUN-2003 (first entry)
XX	DE	Human NOVX polypeptide #94.
XX	KW	Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW	KW	hypertension; congenital heart defect; aortic stenosis; valve disease;
KW	KW	atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW	KW	pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW	KW	tuberculous sclerosis; scleroderma; atherosclerosis; infectious disease;
KW	KW	obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW	KW	Parkinson's disease; immune disorder; haematopoietic disorder;
KW	KW	haemophilia; hypercoagulation; Crohn's disease; cancer.
XX	OS	Homo sapiens.

XX	WO200281498-A2.	
XX	17-OCT-2002.	
XX	03-APR-2002; 2002WO-US010780.	
XX	03-APR-2001; 2001US-0281086P.	
XX	03-APR-2001; 2001US-0281136P.	
XX	05-APR-2001; 2001US-0281863P.	
XX	06-APR-2001; 2001US-0281906P.	
XX	10-APR-2001; 2001US-0282020P.	
XX	10-APR-2001; 2001US-0282930P.	
XX	12-APR-2001; 2001US-0282934P.	
XX	12-APR-2001; 2001US-028312P.	
XX	13-APR-2001; 2001US-0283110P.	
XX	17-APR-2001; 2001US-0284234P.	
XX	19-APR-2001; 2001US-0285325P.	
XX	20-APR-2001; 2001US-0285381P.	
XX	20-APR-2001; 2001US-0285609P.	
XX	23-APR-2001; 2001US-0285748P.	
XX	23-APR-2001; 2001US-028590P.	
XX	24-APR-2001; 2001US-0286068P.	
XX	25-APR-2001; 2001US-0286292P.	
XX	27-APR-2001; 2001US-0287213P.	
XX	02-MAY-2001; 2001US-0288257P.	
XX	29-MAY-2001; 2001US-0294164P.	
XX	30-MAY-2001; 2001US-0294484P.	
XX	18-JUN-2001; 2001US-0298952P.	
XX	19-JUN-2001; 2001US-0299237P.	
XX	19-JUN-2001; 2001US-0299276P.	
XX	12-SEP-2001; 2001US-0318750P.	
XX	25-SEP-2001; 2001US-0324800P.	
XX	25-SEP-2001; 2001US-0324802P.	
XX	27-SEP-2001; 2001US-0325684P.	
XX	17-OCT-2001; 2001US-0330143P.	
XX	14-NOV-2001; 2001US-0332111P.	
XX	14-NOV-2001; 2001US-0332240P.	
XX	14-NOV-2001; 2001US-0332779P.	
XX	21-NOV-2001; 2001US-0332115P.	
XX	04-DEC-2001; 2001US-0337621P.	
XX	03-JAN-2002; 2002US-0345783P.	
XX	16-JAN-2002; 2002US-0350251P.	
XX	02-APR-2002; 2002US-00114270.	
XX	(CURA-) CURAGEN CORP.	
XX	Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;	
PI	Patturajan M, Liu X, Gusev VY, Li L, Vernet CAM, Zerhusen BD;	
PI	Gorman L, Shenoy SG, Pena CEA, Smithson G, Burgess CE, Garlach V;	
PI	Padigaru M, Shimkets RA, Gangolli EA, Taupier RJ, Casman SJ, Ji W;	
PI	Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DU;	
PI	Macdougall JR, Rothenberg MB, Mazur A, Millet I, Peyman JA;	
PI	Ellerman K;	
XX	WPI; 2003-046858/04.	
DR	N-PSDB; ABX72263.	
XX	New isolated NOVX polypeptide useful for treating atherosclerosis,	
PT	metabolic disorders, diabetes, obesity, infectious disease, anorexia,	
PT	neurodegenerative disorders, Alzheimer's disease and cancer.	
XX	Claim 1; Page 292; 665pp; English.	
XX	The invention relates to human polypeptides, termed NOVX, and the	
XX	polynucleotides encoding them. The polypeptides and polynucleotides are	
CC	useful for diagnosing disease, and screening for potential therapeutic	
CC	agents. The sequences are useful for treating metabolic disorders,	
CC	cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic	
CC	stenosis, atrial septal defect (ASD), atrioventricular canal defect,	
CC	ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular	
CC	septal defect (VSD), valve diseases, tuberculous sclerosis, scleroderma,	
CC	atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative	

CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
CC and cancer. Sequences ABU5452-ABU54647 represent human NOVX polypeptides
CC of the invention
XX
SQ Sequence 403 AA;

Query Match 58.4%; Score 2111.5; DB 6; Length 403;
Best Local Similarity 99.0%; Pred. No. 5.6e-157;
Matches 399; Conservative 2; Mismatches 1; Indels 1; Gaps 1;
QY 1 MDKYDVIKAIKOGAFKAYLAKGSDSKHCVIKKEINFEKMPIQKEASKKEVILLERKMH 60
DB 1 MDKYDVIKAIKOGAFKAYLAKGSDSKHCVIKKEINFEKMPIQKEASKKEVILLERKMH 60
QY 61 PNIVAFNFSQNGRLFIWMEYCDGGDLMKRINRQGVLFSEDOILGWFWQISLGLKHH 120
DB 61 PNIVAFNFSQNGRLFIWMEYCDGGDLMKRINRQGVLFSEDOILGWFWQISLGLKHH 120
QY 121 DRKILHRDIKAGNIFLSKNGMVAKLDFGIARVLNNSMELARTCIGTPYLSPEICQKP 180
DB 121 DRKILHRDIKAGNIFLSKNGMVAKLDFGIARVLNNSMELARTCIGTPYLSPEICQKP 180
QY 181 YNNKTDIWSLGCVLVELCTLKHFFEGNNLQOLVLKICQAHFAPISPGFSRELHSLISOLF 240
DB 181 YNNKTDIWSLGCVLVELCTLKHFFEGNNLQOLVLKICQAHFAPISPGFSRELHSLISOLF 240
QY 241 QVSPDRSINSILKRPFLNLIPLYLTPV-IOEFSHMLICRAGAPASHAGKVVQKC 299
DB 241 QVSPDRSINSILKRPFLNLIPLYLTPV-IOEFSHMLICRAGAPASHAGKVVQKC 300
QY 300 KIQKVFQKCPKPPRSRISVPKRNAILHRNWRPPAGAKARSIMIERPKIAVCGHYD 359
DB 301 KIQKVFQKCPKPPRSRISVPKRNAILHRNWRPPAGAKARSIMIERPKIAVCGHYD 360
QY 360 YYYAQLDMLRRRAHKPSVHPDQENTGVEDYQETRHGSPSQ 402
DB 361 YYYAQLDMLRRRAHKPSVHPDQENTGVEDYQETRHGSPSQ 403

RESULT 7
ABU54634
ID ABU54634 standard; protein; 399 AA.
XX
AC ABU54634;
XX
DT 03-JUN-2003 (first entry)
DE Human NOVX polypeptide #93.
XX

Human; NOVX; metabolic disorder; cardiomyopathy; diabetes; ASD;
KW hypertension; congenital heart defect; aortic stenosis; valve disease;
KW atrial septal defect; atrioventricular canal defect; ductus arteriosus;
KW pulmonary stenosis; subaortic stenosis; ventricular septal defect; VSD;
KW tuberosus sclerosis; scleroderma; atherosclerosis; infectious disease;
KW obesity; anorexia; neurodegenerative disorder; Alzheimer's disease;
KW Parkinson's disease; immune disorder; haematopoietic disorder;
KW haemophilia; hypercoagulation; Crohn's disease; cancer.
XX
OS Homo sapiens.
XX
PN WC200281498-A2.
XX
PD 17-OCT-2002.
XX
PF 03-APR-2002; 2002MO-US010780.
XX
PR 03-APR-2001; 2001US-0281086P.
PR 03-APR-2001; 2001US-0281136P.
PR 05-APR-2001; 2001US-0281863P.
PR 05-APR-2001; 2001US-0281906P.
PR 06-APR-2001; 2001US-0282028P.
PR 10-APR-2001; 2001US-0282930P.

PR 10-APR-2001; 2001US-0282934P.
PR 12-APR-2001; 2001US-0283512P.
PR 13-APR-2001; 2001US-0283710P.
PR 17-APR-2001; 2001US-0284234P.
PR 19-APR-2001; 2001US-0285325P.
PR 20-APR-2001; 2001US-0285381P.
PR 20-APR-2001; 2001US-0285609P.
PR 23-APR-2001; 2001US-0285748P.
PR 23-APR-2001; 2001US-0285890P.
PR 24-APR-2001; 2001US-0286068P.
PR 25-APR-2001; 2001US-0286292P.
PR 27-APR-2001; 2001US-0287213P.
PR 02-MAY-2001; 2001US-0288257P.
PR 29-MAY-2001; 2001US-0294164P.
PR 30-MAY-2001; 2001US-0294484P.
PR 18-JUN-2001; 2001US-0298952P.
PR 19-JUN-2001; 2001US-0299237P.
PR 19-JUN-2001; 2001US-0299276P.
PR 12-SEP-2001; 2001US-0318750P.
PR 25-SEP-2001; 2001US-0324800P.
PR 25-SEP-2001; 2001US-0324802P.
PR 27-SEP-2001; 2001US-0325684P.
PR 17-OCT-2001; 2001US-0330143P.
PR 14-NOV-2001; 2001US-0332131P.
PR 14-NOV-2001; 2001US-0332240P.
PR 14-NOV-2001; 2001US-0332779P.
PR 21-NOV-2001; 2001US-0332115P.
PR 04-DEC-2001; 2001US-0337621P.
PR 03-JAN-2002; 2002US-0345783P.
PR 16-JAN-2002; 2002US-0350251P.
PR 02-APR-2002; 2002US-00114270.
XX
XX (CURA-) CURAGEN CORP.
XX
PI Guo X, Kekuda R, Miller CE, Malyankar UM, Spytek KA;
PI Patturajan M, Liu X, Gusev VY, Li L, Vernet CAM, Zerkhusen BD;
PI Gorman L, Shency SG, Pena CE, Smithson G, Burgess CE, Gerlach V;
PI Padigaru M, Shimkets RA, Gangoli EA, Taupier RJ, Casman SJ, Ji W;
PI Anderson DW, Leite MW, Rastelli L, Edinger SR, Stone DJ;
PI Maccougall JR, Rothenberg ME, Mazur A, Millet I, Peyman JA;
PI Ellerman K;
XX
XX WPI; 2003-046858/04.
DR N-PSDB; ABX72262.
XX
XX New isolated NOVX polypeptide useful for treating atherosclerosis,
PT metabolic disorders, diabetes, obesity, infectious disease, anorexia,
PT neurodegenerative disorders, Alzheimer's disease and cancer.
XX
XX Claim 1; Page 291; 666pp; English.
XX
XX The invention relates to human polypeptides, termed NOVX, and the
CC polynucleotides encoding them. The polypeptides and polynucleotides are
CC useful for diagnosing disease, and screening for potential therapeutic
CC agents. The sequences are useful for treating metabolic disorders,
CC cardiomyopathy, diabetes, hypertension, congenital heart defects, aortic
CC stenosis, atrial septal defect (ASD), atrioventricular canal defect,
CC ductus arteriosus, pulmonary stenosis, subaortic stenosis, ventricular
CC septal defect (VSD), valve diseases, tuberosus sclerosis, scleroderma,
CC atherosclerosis, obesity, infectious disease, anorexia, neurodegenerative
CC disorders, Alzheimer's disease, Parkinson's disease, immune disorders,
CC haematopoietic disorders, haemophilia, hypercoagulation, Crohn's disease
CC and cancer. Sequences ABU5452-ABU54647 represent human NOVX polypeptides
CC of the invention
XX
XX Sequence 399 AA;
SQ
Query Match 58.3%; Score 2106.5; DB 6; Length 399;
Best Local Similarity 99.3%; Pred. No. 1.4e-156;
Matches 399; Conservative 0; Mismatches 0; Indels 3; Gaps 1;
QY 1 MDKYDVIKAIKOGAFKAYLAKGSDSKHCVIKKEINFEKMPIQKEASKKEVILLERKMH 60

Db 1 MDKVDYKAIGQAFKAYLAKGSDSKHCVIKEINFEK---QEKEASKKEVILLKMKH 57
QY 61 PNIVAFNSFQNGRLFIWMEYCDGDLMKRINRQGVLFSEDOILGHFVQISLGLKH 120
Db 58 PNIVAFNSFQNGRLFIWMEYCDGDLMKRINRQGVLFSEDOILGHFVQISLGLKH 117
QY 121 DRKILHRDIKAONIPLSKNGMWAKLGDGFIARVLNNSMELARTCTGTPTYYLSPEICQNK 180
Db 118 DRKILHRDIKAONIPLSKNGMWAKLGDGFIARVLNNSMELARTCTGTPTYYLSPEICQNK 177
QY 181 YNNKTDIWSLGCVLVELCTLKHPPFEGNLLQVLVLCIQAHPAPISPGFSRSLHSLSOLF 240
Db 178 YNNKTDIWSLGCVLVELCTLKHPPFEGNLLQVLVLCIQAHPAPISPGFSRSLHSLSOLF 237
QY 241 QVSPDRPSINSILKRPFLNLIIPKYLTPVIEQEEFSLHMLICRAGAPASRHAGKVQCK 300
Db 238 QVSPDRPSINSILKRPFLNLIIPKYLTPVIEQEEFSLHMLICRAGAPASRHAGKVQCK 297
QY 301 IQKVFQKCPSPRSISVPIKRNALHNRNEPPAGAKASIKMIERPKIAAVCGHYD 360
Db 298 IQKVFQKCPSPRSISVPIKRNALHNRNEPPAGAKASIKMIERPKIAAVCGHYD 357
QY 361 YYAQLDMLRRRAHKPSYHPIPOENTGVEDYGOETRHGSPSQ 402
Db 358 YYAQLDMLRRRAHKPSYHPIPOENTGVEDYGOETRHGSPSQ 399

RESULT 8
ABP96073 standard; protein; 544 AA.
AC ABP96073;
XX
DT 07-MAY-2003 (first entry)
DE Human protein kinase SEQ ID NO:52.
XX Human; protein kinase; enzyme; antiasthmatic; antiinflammatory;
KW antidiabetic; antiparkinsonian; antimigraine; cardiant; cytostatic;
KW immunosuppressive; vulnarary; gene therapy; COPD; asthma; migraine;
KW chronic obstructive pulmonary disease; non-insulin dependent diabetes;
KW Parkinson's disease; myocardial infarction; inflammatory bowel disease;
KW autoimmune disorder; allograft rejection; graft versus host disease;
KW cancer; leukaemia; wound granulation.
XX Homo sapiens.
XX WO2003000901-A2.
XX 03-JAN-2003.
XX 24-JUN-2002; 2002WO-IB002358.
XX 26-JUN-2001; 2001US-0301098P.
XX 06-NOV-2001; 2001US-0332870P.
XX (DECO-) DECODE GENETICS EHF.
XX Martinez RAM, Sigurdsson GT;
PI WPI; 2003-201429/19.
DR N-PSDB; ABZ77151.
XX New protein kinase genes and polypeptides, useful for diagnosing diseases
PT associated with a protein kinase, or in gene therapy for treating e.g.
PT Parkinson's disease, migraine, myocardial infarction, allograft rejection
PT or cancers.
XX Claim 9; Page 75-76; 258pp; English.
XX ABZ77126 to ABZ77165 encode the human protein kinases given in ABP96048
CC to ABP96087. The protein kinases have antiasthmatic, antiinflammatory,
CC antidiabetic, antiparkinsonian, antimigraine, cardiant, cytostatic,

CC immunosuppressive and vulnarary activities, and can be used in gene
CC therapy. A protein kinase therapeutic agent from the present invention,
CC particularly a protein kinase gene agonist or antagonist, can be used for
CC treating a disease or condition associated with a protein kinase in an
CC individual. These diseases include chronic obstructive pulmonary diseases
CC (COPD), asthma, non-insulin dependent diabetes, Parkinson's disease,
CC migraine, myocardial infarction, inflammatory bowel disease, autoimmune
CC disorders (e.g. allograft rejection or graft vs. host disease), cancers
CC (e.g. leukaemias) or wound granulation
XX Sequence 544 AA;
SQ Query Match 30.9%; Score 1118.5; DB 6; Length 544;
Best Local Similarity 39.8%; Pred. No. 7.2e-79;
Matches 290; Conservative 38; Mismatches 112; Indels 289; Gaps 21;
QY 10 IGCAF--GKAYLAKGSDSKHCVIKEINFEKMPIQEKEASKKEVILLKMKHNPVAPF 67
Db 16 VGLGVFLRGVAYAKRERDGEWLL-----SGAVKXHALIKNMAD----KYL 58
QY 68 NSFQ-----ENGRLFIVMEYCDGDLMKRINRQGVLFSEDOILGHFVQISLGLKH 119
Db 59 KSFQHQMIPLTENGRLFIWMEYCDGDLMKRINRQGVLFSEDOILGHFVQISLGLKH 118
QY 120 HDRKILHRDIKAONIPLSKNGMWAKLGDGFIARVLNNSMELARTCTGTPTYYLSPEICQNK 179
Db 119 HDRKILHRDIKAONIPLSKNGMWAKLGDGFIARVLNNSMELARTCTGTPTYYLSPEICQNK 178
QY 180 PYNNKTDIWSLGCVLVELCTLKHPPFEGNLLQVLVLCIQAHPAPISPGFSRSLHSLSOLF 239
Db 179 PYNNKTDIWSLGCVLVELCTLKHPPFEGNLLQVLVLCIQAHPAPISPGFSRSLHSLSOLF 238
QY 240 QVSPDRPSINSILKRPFLNLIIPKYLTPVIEQEEFSLHMLICRAGAPASRHAGKVQCK 299
Db 239 QVSPDRPSINSILKRPFLNLIIPKYLTPVIEQEEFSLHMLICRAGAPASRHAGKVQCK 298
QY 300 KIQKVFQKCPSPRSISVPIKRNALHNRNEPPAGAKASIKMIERPKIAAVCGHYD 359
Db 271 -----VSSKGYLRQNLSAVEL-----ILKNLLK-APHWKFILSPKKVSR 331
QY 469 -----QLEIRIQVYHNDMKEIRKMGREPEENSKISH-----KTYLVKX 507
Db 332 VGTDHCFRRASHEIASLVQAQPHM-AGEASQSQKAKEEQSHALHEGROESLWELLFTKP 390
QY 508 SNL----PVHQDASEGEAPVDIEKDLQVRLQNTKESKN-----PEQKYKAKGVKFEIN 559
Db 391 SLLKAAVSHDCATVPQGIQ-----SKTRSQKNKALLTSQTRRGKGAAPH--- 436
QY 560 LDKCISDENILOEEBAMDIPNETLTFEDGMKFKEVCVKEHGDYTDKAFELKHCPEAGFS 619
Db 437 -----ISDDG---RPGRDAPH-----FLDGM-----AAGQR 459
QY 620 TQTVAAVGNR-----ROWDGGAPQILLQWMAVADITCTCPTGPDSESVLSV- 666
Db 460 RSSLSRLGSAEGLTSTQMGGAETLL-----TSQTGLRQENCLNLGGGCGSEP 509
QY 667 RQEGKTKDP 675
Db 510 RSSGERGHP 518

RESULT 9
AAO26613
ID AAO26613 standard; protein; 774 AA.
XX

PD 03-JAN-2003.
XX 24-JUN-2002; 2002WO-BF006948.
PF 25-JUN-2001; 2001US-0300068P.
PR 07-DEC-2001; 2001US-0336704P.
XX (FARB) BAYER AG.
PA Xiao Y;
XX WPI; 2003-184051/18.
PI
XX New polynucleotide encoding a NEK-like serine/threonine kinase
PT polypeptide useful for treating diseases associated with kinase
PT dysfunction, e.g. cardiovascular disorders, cancer such as colon cancer,
PT diabetes and CNS disorders.
XX
XX Disclosure; Fig 3; 149pp; English.
PS The invention relates to a newly isolated polynucleotide encoding an NEK-
CC like serine/threonine protein kinase. The activity of the polynucleotide
CC and polypeptide of the invention may be described as cytostatic,
CC cardiant, antiinflammatory, nontropic and neuroprotective. The expression
CC vector and reagent of the invention are useful for the preparation of a
CC medicament for modulating the activity of an NEK-like serine/threonine
CC kinase in a disease, such as cancer (e.g. colon cancer), cardiovascular
CC disorder, diabetes, COPD or CNS disorder. The polypeptides may also be
CC used to identify compounds which may act as activators or inhibitors at
CC the enzyme's active site, to raise specific antibodies which can block
CC the enzyme and effectively reduce its activity, as a bait protein in a
CC two-hybrid or three-hybrid assay to identify other proteins which bind to
CC or interact with the human NEK-like serine/threonine kinase polypeptide
CC and modulate its activity, and for the immunisation of mammals. The
CC current sequence represents an NEK-like serine/threonine protein kinase
CC designated NEK1_mouse
XX
XX Sequence 774 AA;
SQ
Query Match 30.6%; Score 1106.5; DB 6; Length 774;
Best Local Similarity 35.2%; Pred. No. 1e-77;
Matches 276; Conservative 121; Mismatches 207; Indels 181; Gaps 23;
QY 1 MDKYDVIKAGGARGKAYLAGKSDSKHCYIKENIFKPMIOEAKSKVEILLKMKH 60
DB 1 MEKYVRLQKIGESGKAVLVKSTEDGRHYVFIKEINISMSDKERQESREVRVLANMKH 60
QY 61 PNIVAFNSFQNGRLFTVMEYCDGGLMKRINRGVLFSDQILGMFVQISLGLKH 120
DB 61 PNIVQKSPFENGSLYVMDYCEGDLFKRNAQKALFQEDQLLDHFWICLAKH 120
QY 121 DRKILHRDIKQNIPLSKNGWAKLGDGFIARVLNNSMELARTCTGTPTYYLSPEICQNP 180
DB 121 DRKILHRDIKQNIPLSKNGWAKLGDGFIARVLNNSMELARTCTGTPTYYLSPEICQNP 179
QY 181 YNKTDIWSLGCYVLYELCTLKHPPFGNNLQQLVLIQAHFAPISPGFSRHLISLGLF 240
DB 180 YNKSDDIWSLGCYVLYELCTLKHPPFGNNLQQLVLIQAHFAPISPGFSRHLISLGLF 239
QY 241 QVSPDRPSINSILKRPFLNLIPIKYLPEVITQEEFESHMLICRAG----- 285
DB 240 KENPRDRPSVNSILEKGFIAKRIEFLSPQIAEFLKTLKSPFPVSPHYSDIRLSLQ 299
QY 286 -----APASHAGKVQCKIQKRFQKCPSPSRIS--VPIKENAILHENEWRP 333
DB 300 SSFVPAQKITPAKYGVLPTTKYKGYDKLLEKPPPKHQAHPVKK---MNSGEERK 356
QY 334 PAG--AQKARSIMTER----- 348
DB 357 KMSEAAKRRLEFIEKEKKQDKQIRFLKAEQMKPQEKQRLERINRAREQGNVLRAGG 416
QY 349 -----PKIAVCHYDYVYQAOLDMLRR-RAH-----KPSYHP--IQE 383

Db 417 SGEVKASFFGIGGAVSPSPGSGYEHYHAIQDQRLAEADNEARWKGIGYGRWLPER 476
QY 384 NTG---VEDYQGETRHGSPSPQWPAEYLQSKFPAQYKLVKVE-----KQGLRPPSS 431
Db 477 QKGLAVERANQ-----VEBFLOKREAMONKARAEHGHHVYVYLAELRQIRLQ--- 522
QY 432 AEPNYNORQLR-----SNGEPRFOELPRKNEMKEQEVYKQLEIEIRQQYHN 479
Db 523 ---NFNERQIKAKLGENKEADGTGQEQAT--ETDMR---LKMESLKAQTWARAVLK 575
QY 480 DMKEIRKMGREPEPENSKISHKTYLVKKNLPHVQDASE-GEAPVODIEKDKMLQNT 538
Db 576 EQLERKKEAYEREKKVWEHLVARVKSVDPLPLELLETGGSP----- 619
QY 539 KESKNPEQYKAKGVKFEINLDCISDENILQEEBAMDPNETLT-----FSDGMK 590
Db 620 --SKQVKPVISTALKVEGLDGLTD-----TQEEEMKSNASISKREILRLNENLK 673
QY 591 FKEYECVKEHGYDITKAPEKLHCPCEAGFSTQTVAAVGNRRQWDGGAPOTLQMMVADIT 650
Db 674 AQDEKEKQH--HSGCETVGHKDEREYETEN-AISSDRKKWEMGG-QLVPLDAVTLDT 729
QY 651 STCPT 655
Db 730 SFSAT 734
RESULT 11
AAO16441
ID AAO16441 standard; protein; 774 AA.
XX AAO16441;
AC AAO16441;
DT 10-APR-2003 (first entry)
XX Human serine/threonine protein kinase NEK1-related protein.
DE
XX Human; gene therapy; serine/threonine protein kinase; NEK1; cancer;
XX colon cancer; cardiovascular disorder; congestive heart failure;
XX central nervous system disorder; chronic obstructive pulmonary disease;
XX CNS disorder; diabetes; myocardial infarction; ischaemic heart disease;
XX arrhythmia; hypertensive; Alzheimer's disease; Parkinson's disease;
XX peripheral pain; chronic pain.
XX
OS Homo sapiens.
XX
XX WO2003000873-A2.
XX
XX 03-JAN-2003.
XX
XX 21-JUN-2002; 2002WO-BF006879.
XX
XX 25-JUN-2001; 2001US-0300071P.
XX 16-NOV-2001; 2001US-0331447P.
XX 07-DEC-2001; 2001US-0336693P.
XX
XX (FARB) BAYER AG.
XX
XX Xiao Y;
XX
XX WPI; 2003-201424/19.
XX
XX New serine/threonine protein kinase NEK1 gene and protein, useful for
XX identifying modulators of serine/threonine protein kinase NEK1 activity,
XX and in gene therapy for treating cancer, diabetes, heart failure or
XX Alzheimer's disease.
XX
XX Disclosure; Fig 3; 156pp; English.
XX
XX The invention comprises the amino acid and coding sequence of the human
XX serine/threonine protein kinase NEK1. The DNA and protein sequences of
XX the invention are useful for modulating the activity of serine/threonine
XX kinase NEK1 in a disease, such as: cancer (particularly colon cancer);

CC cardiovascular disorders; central nervous system (CNS) disorders;
 CC diabetes; and chronic obstructive pulmonary disease. In particular the
 CC DNA and protein sequences of the invention are useful for treating:
 CC congestive heart failure; myocardial infarction; ischaemic heart disease;
 CC arrhythmia; hypertensive; Alzheimer's disease; Parkinson's disease; and
 CC peripheral or chronic pain. The present amino acid sequence represents a
 CC human serine/threonine protein kinase NEK1-related protein
 XX
 SQ Sequence 774 AA;
 Query Match 30.6%; Score 1106.5; DB 6; Length 774;
 Best Local Similarity 35.2%; Pred. No. 1e-77;
 Matches 276; Conservative 121; Mismatches 207; Indels 181; Gaps 23;
 QY 1 MDKYDVIKAGGAFKAYLAGKSDSKHCVIKINFEKMPQKEASKEVILEKMKH 60
 DB 1 MEKYVRLQKIGEGSGKAVLVKSTEDGRHYVIKINISRMDSKQESREAVLANMKH 60
 QY 61 PNIVAFFNSFOENGRFLFVMEYCDGDLMKRINRQGVLFSEDOILGFWQISLGLKH 120
 DB 61 PNIVQKESFEENGSLYVMDYCEGDLFKRINAKGALFOEDQILDWFVQICLAKH 120
 QY 121 DRKILHRDIKQNIPLSKNGVMVAKLGDGFIARVLNNSMELARTCIGTPYLSPICONKP 180
 DB 121 DRKILHRDIKQNIPLSKNGVMVAKLGDGFIARVLNNSMELARTCIGTPYLSPICONKP 179
 QY 181 YNNKTDIWSLGLVLYELCTLKHPPFEGNNLQQLVKIQAHFAPISPGFSRELHSLISQLF 240
 DB 180 YNNKSDIWLALGCVLYELCTLKHAFEGAGNKNLVLIKISGSPVPSHYSDRLSLLQLP 239
 QY 241 QVSPDRPSINSILKRPFLNLPKYLTPVIOSEFSLMLICRAG----- 285
 DB 240 KENPRDRPSVNSILEKGFIAKRIEFLSPQIAEFCILKTLKFGPQPLPKRPSAQGV 299
 QY 286 -----APASHACKVQVKCKIQVRFQGGKCPSPRIS--VPIKRNAILHRENWRP 333
 DB 300 SFVPAQKITPAKRYGVLTKYKDKLEKPPKPKHQAQHPVK-----MNSGEERK 356
 QY 334 PAG--AQKARSIFMIER----- 348
 DB 357 KMSEEAARLEFTEKEKKQKQDIKFLKABQMKRQKQRLERINRAREQGRNVLRAAG 416
 QY 349 -----PKIAVCGHYDYVYQAQLMLR-RAH-----KSYHP--LPQE 383
 DB 417 SGVKAAPFGIGGAVSPSPSPGQYEHYHAFIDQMORLRAEDNEARWKGIGYGRWLPER 476
 QY 384 NTG---VEDYQETRHGSPSPQWPAEYLQKFEAQYKLVKVE-----KQLGLRPS 431
 DB 477 QKGLHAVERANQ-----VEEFLQKREAVQNKARAGHVVLARLQIRLQ--- 522
 QY 432 AEPNYNQOEIR-----SNGEPRFQELPRFNEMKQETVWQLBEIRIQOYHN 479
 DB 523 ---NFERQKIKALRGENKADGTGQEAT--EETDMR---LKKVBSLKQATNARAAVLK 575
 QY 480 DMKEIRKMGREPEENSKIHKTVLKKSNLPVHODASE--GEAPVDIEKDLKOMRLQNT 538
 DB 576 EQLERKREAVEREKVVWEHLVARVKSVDPLLELLETGGSP----- 619
 QY 539 KESKNPEQYKAKGVKFEINLDCISDENILQISEEAMIDPNEYLT-----FEGMK 590
 DB 620 --SKQVFPVSVTSALKEVGLDGLTD---TOEEEMKSNLSAISKREILRLNENLK 673
 QY 591 PKEYECVHEGDDYTDKAFKELHCPAGSPSTQVAVGNRRQWDGAPQTLQOMVADIT 650
 DB 674 AQEDEKEQH--HSGSCSTVGHKBEREYETEN-AISDRKKWENG--QLVPLDVAITLD 729
 QY 651 STCPT 655
 DB 730 SFSAT 734
 RESULT 12.
 ABP97691

ID ABP97691 standard; protein; 774 AA.
 XX ABP97691;
 XX
 DT 16-MAY-2003 (first entry)
 XX
 DE Amino acid sequence of human protein sharing some identity with NEK3.
 XX
 XX Human; serine threonine protein kinase; NEK3; chromosome 13q14.3; cancer;
 KW colon cancer; central nervous system disorder; mood disorder;
 KW anxiety disorder; Parkinson's disease; Alzheimer's disease;
 KW cardiovascular disorder; myocardial infarction; hypertension; arrhythmia;
 KW embolism; acute arterial thrombosis; ischaemic disease; angina pectoris;
 KW chronic obstructive pulmonary disease; diabetes; pain.
 XX
 OS Homo sapiens.
 XX
 FN WO2003000874-A2.
 XX
 PD 03-JAN-2003.
 XX
 PF 25-JUN-2002; 2002WO-EP006993.
 XX
 PR 25-JUN-2001; 2001US-0300067P.
 PR 13-NOV-2001; 2001US-0331225P.
 PR 04-DEC-2001; 2001US-0334952P.
 XX
 PA (FARB) BAYER AG.
 XX
 PI Xiao Y;
 XX
 DR WPI; 2003-221415/21.
 XX
 PT New human serine/threonine protein kinase NEK3 polypeptide and
 PT polynucleotide, useful in preventing, ameliorating, or treating diseases
 PT associated with serine/threonine protein kinase NEK3 dysfunction such as
 PT cancer or diabetes.
 XX
 PS Disclosure; Fig 6; 14pp; English.
 XX
 CC The present sequence represents a human polypeptide, sharing some
 CC identity with a human serine/threonine protein kinase NEK3 polypeptide.
 CC The gene is located on chromosome 13q14.3. The serine/threonine protein
 CC kinase NEK3 polypeptide and polynucleotide are useful in preventing,
 CC ameliorating, or treating diseases associated with serine/threonine
 CC protein kinase NEK3 dysfunction such as cancer, particularly colon
 CC cancer, central nervous system disorders (e.g. mood disorders, anxiety
 CC disorders, Parkinson's disease or Alzheimer's disease), cardiovascular
 CC disorders (e.g. myocardial infarction, hypertension, arrhythmias,
 CC embolism, acute arterial thrombosis, ischaemic diseases or angina
 CC pectoris), chronic obstructive pulmonary disease, or diabetes. These can
 CC also be used to treat pain associated with the above disorders. The
 CC serine/threonine protein kinase NEK3 polypeptide is also useful in
 CC diagnostic assays or in genetic testing
 XX
 SQ Sequence 774 AA;
 Query Match 30.6%; Score 1106.5; DB 6; Length 774;
 Best Local Similarity 35.2%; Pred. No. 1e-77;
 Matches 276; Conservative 121; Mismatches 207; Indels 181; Gaps 23;
 QY 1 MDKYDVIKAGGAFKAYLAGKSDSKHCVIKINFEKMPQKEASKEVILEKMKH 60
 DB 1 MEKYVRLQKIGEGSGKAVLVKSTEDGRHYVIKINISRMDSKQESREAVLANMKH 60
 QY 61 PNIVAFFNSFOENGRFLFVMEYCDGDLMKRINRQGVLFSEDOILGFWQISLGLKH 120
 DB 61 PNIVQKESFEENGSLYVMDYCEGDLFKRINAKGALFOEDQILDWFVQICLAKH 120
 QY 121 DRKILHRDIKQNIPLSKNGVMVAKLGDGFIARVLNNSMELARTCIGTPYLSPICONKP 180
 DB 121 DRKILHRDIKQNIPLSKNGVMVAKLGDGFIARVLNNSMELARTCIGTPYLSPICONKP 179

QY 181 YNNKTDINSLGCVLYELCTLKHPFPGNNLQOLVLKICQAHFAPISPGFSLHLSLSOLF 240
Db 180 YNNKSDIHWALGCVLYELCTLKHPFAGNNKLVLIISGSPFPPVPHYSYDLRSLSOLF 239
QY 241 QVSPDRPSINSILKRPFLNIPKYLTPFVIOEFSHMLICRAG----- 285
Db 240 KRNPRDRPSVNSILEKGFIAKRIEFLGSPQIAEFLCKLTKSGFQPLPGKRPASGGGV 299
QY 286 -----APASHAGKVVQCKIKQVRQKCPSPRSIS--VPIKRNAILHNEWRP 333
Db 300 SFVPAQKITKPAKYGIPLVTKYKYGDKLLEKPPKQAHQIPVKK--MNSGEERK 356
QY 334 PAG--AQKARSIMIER----- 348
Db 357 KMSEAAKRRLEFIEKEKKQDKQIRFLKAEQMKQEKQRLERINRAREQGNVLRAGG 416
QY 349 -----PKIAVCGHYDYVYAQLDMLRR-RAH-----KPSYHP--IPQE 383
Db 417 SGFVKASFFGIGAVSPSPSPRGQYEHYHAIFDQMLRAEDNEARWKGIYGRWLPER 476
QY 384 NTG---VEDYGGQTRHGPPSPQWPAEYLQKPEAQYKLVKVE-----KQLGLRPS 431
Db 477 QKGLHVAERANQ-----VEEFLQKREANQKARAEHVVYLARLQRLQ--- 522
QY 432 ABPNVNRQELR-----SNGEPRPQELPFRKNEMKEOYVWKQLEETROQVHN 479
Db 523 ---NFNERQOIKAKLRGENKEADGKGQEAT--EETDMR---LKKMESLKQAQTNARAALV 575
QY 480 DMKETRKMGPPPEENSKISHYTKLVKSNLPHVDASE-GEAPVDQIDEKDKQMLQNT 538
Db 576 EQLERKREAYEREKWEHVLVAVKSDVPLPLELLETGSP----- 619
QY 539 KESKNPEQYKAKGKVFENLDKICSDENILQEEANDIPNETL-----FEDGMK 590
Db 620 --SKQVKPVISVTSALKVEVGLDGLTD---TQEEEMKSNASISKREILRLNENLK 673
QY 591 FKEYECVKEHGDYTDKAEKELHCPAGSPSTQTVAAVGNRQWDGAPQTLQMAVADIT 650
Db 674 AQEDEKEKQH--HSGSCETVGHKDBREYETEN-AISDRKKWEMGG-QLVPLDAVTLDT 729
QY 651 STCPT 655
Db 730 SPSAT 734
RESULT 13
ID AAM39210 standard; protein; 1242 AA.
AC AAM39210;
XX
DT 22-OCT-2001 (first entry)
DE Human polypeptide SEQ ID NO 2355.
KW Human; nootropic; immunosuppressant; cytostatic; gene therapy; cancer;
KW peripheral nervous system; neuropathy; central nervous system; CNS;
KW Alzheimer's; Parkinson's disease; Huntington's disease; haemostatic;
KW amyotrophic lateral sclerosis; Shy-Drager Syndrome; chemotactic;
KW chemokinetic; thrombolytic; drug screening; arthritis; inflammation;
KW leukaemia.
OS Homo sapiens.
XX
XX WO200153112-A1.
XX
XX 26-JUL-2001.
XX
XX 26-DEC-2000; 2000WO-US034263.
XX
XX 23-DEC-1999; 99US-00471275.
PR 21-JAN-2000; 2000US-00488725.
PR 25-APR-2000; 2000US-00552317.
PR

PR 20-JUN-2000; 2000US-00598042.
PR 19-JUL-2000; 2000US-00620312.
PR 03-AUG-2000; 2000US-00653450.
PR 14-SEP-2000; 2000US-00662191.
PR 19-OCT-2000; 2000US-00693036.
PR 29-NOV-2000; 2000US-00727344.
PA (HYSE-) HYSEQ INC.
XX Tang YT, Liu C, Asundi V, Chen R, Ma Y, Qian XB, Ren F, Wang D;
PI Wang J, Wang Z, Wehrman T, Xu C, Xue AJ, Yang Y, Zhang J, Zhao QA;
PI Zhou P, Goodrich R, Drmanac RT;
XX WPI; 2001-442253/47.
DR N-PSDB; AAI58366.
XX
XX Novel nucleic acids and polypeptides, useful for treating disorders such
XX as central nervous system injuries.
PS Example 4; SEQ ID NO 2355; 10078pp; English.
XX
XX The invention relates to human nucleic acids (AAI57798-AAI61369) and the
XX encoded polypeptides (AAM38642-AAM42213) with nootropic,
XX immunosuppressant and cytostatic activity. The polynucleotides are useful
XX in gene therapy. A composition containing a polypeptide or polynucleotide
XX of the invention may be used to treat diseases of the peripheral nervous
XX system, such as peripheral nervous injuries, peripheral neuropathy and
XX localised neuropathies and central nervous system diseases, such as
XX Alzheimer's, Parkinson's disease, Huntington's disease, amyotrophic
XX lateral sclerosis, and Shy-Drager Syndrome. Other uses include the
XX utilisation of the activities such as: immune system suppression,
XX Activin/inhibin activity, chemotactic/chemokinetic activity, haemostatic
XX and thrombolytic activity, cancer diagnosis and therapy, drug screening,
XX assays for receptor activity, arthritis and inflammation, leukaemias and
XX C.N.S disorders. Note: The sequence data for this patent did not form
XX part of the printed specification
SQ Sequence 1242 AA;
Query Match 30.6%; Score 1105.5; DB 4; Length 1242;
Best Local Similarity 33.7%; Pred. No. 2.4e-77;
Matches 283; Conservative 121; Mismatches 208; Indels 229; Gaps 22;
QY 1 MDXYDVKAICGAGFAGKAYLAKGSKSKVCVKEINFEKMPIOEKEASKEVILLKMKH 60
Db 1 MEKYVRLQKIGEGFGKAILVKTEDGRQYVKEINISRMSSKERESREAVLANMKH 60
QY 61 PNIVAFNPSFQENGRFLFVMEYCDGDLMKRINRQGVLFSEDIILGWFIQSLGLKHH 120
Db 61 PNIVQYRESPEENGSLYVMDYCEGDLPRINAQGVLFQEDQILDWFIQICLAKHHV 120
QY 121 DRKILHRDIKAQNIPLSKNGWAKLGDGFIARVNNMELARTICITPYLSPICQKPK 180
Db 121 DRKILHRDIKSNIFLTQDGTV-QLGDFGIARVNSTVELARTICITPYLSPICENKP 179
QY 181 YNNKTDIWSLGGCVLYELCTLKHPFEGNNLQOLVLKICQAHFAPISPGFSLHLSLSOLF 240
Db 180 YNNKSDIHWALGCVLYELCTLKHPFAGNNKLVLIISGSPFPPVPHYSYDLRSLSOLF 239
QY 241 QVSPDRPSINSILKRPFLNIPKYLTPFVIOEFSHMLICRAG---PASHAG--- 293
Db 240 KRNPRDRPSVNSILEKGFIAKRIEFLGSPQIAEFLCKLTKSGFQPLPGKRPASGGVS 299
QY 294 -----KVQCKIKQVRQKCPSPRSISVPIKRNAI 325
Db 300 ISVNPQKITKPAKYGIPLVTKYKYGDKLLEKPPKQAHQIPVKK--WRPPAGAQA 340
QY 326 -----LHRNE----- 358
Db 359 SEEAARKRLEFIEKEKKQDKQIISLMKAEQMKQEKERLERINRAREQGNVLRAGG 418
QY 341 RSIMK-----IERPKIAAVCGHYDYVYAQLDMLRRRAKPSYHPIQENTGVDEYG-- 391
PR

Db 419 GEVKAPFLGSGGTIAPSPSSRGQYEHYHAIFDQMQQRAEDN-----EAKWKREIYGRG 473
 QY 392 ----QETHRGPSPSQWPAEYLQKFEAQYKLVKVEKQLGL-----RPSSAEPNPN 437
 Db 474 LPERQKQLAVERAKQVEEFLQKREAMQKARAEHGMILQNLAAVMYGGPSS-----527
 QY 438 QRQELRSNGEPRFOELPRKNEKQEWKQLEIRQYHNDKMEIRKQM-----488
 Db 528 -----SRGKPRNKE-----EVLARLQIRLQNFNERQQLAKLRGKKEANH 572
 QY 489 --GREPEENKISHKTYLVKNSLPVHQDASEGEAPVDIEKDLQKMLQNTKESKNEPQ 546
 Db 573 SEGQSGSEADMRK-----KIESLKAHANA---RAAV--LKEQLERKKEAYEREKVVWE 623
 QY 547 KYKAKKGVKF-----EINLQKISDENILOEE- 573
 Db 624 EHLVAKGVKSDVSPPLQGHETGGSPSQQMRSVISVTSALKKEGVDSLSLDTRETSEEM 683
 QY 574 -----EAMDIPNETLTFEDGMKFKE-----YECVKEHGYDT- 605
 Db 684 QKTNNAISSKREILRLNENLKAQDEKQKQNLSDTFEINVHEDAKHEKEKSVSDRKK 743
 QY 606 -KAFELKHP-----EAGFSTQTVAAVGN-----PROWDGGAPQTLQMAVAD 648
 Db 744 WEAGQQLVPLDELTLDT7SFSTTERHTVGEVILKGPNGSPRAWGKSPTDSVLKILGEAE 803
 QY 649 I 649
 Db 804 L 804
 RESULT 14
 AAO16440
 ID AAO16440 standard; protein; 1242 AA.
 XX AC AAO16440;
 DT 10-APR-2003 (first entry)
 DE Human serine/threonine protein kinase NEK1.
 KW Human; gene therapy; serine/threonine protein kinase; NEK1; cancer;
 KW colon cancer; cardiovascular disorder; congestive heart failure;
 KW central nervous system disorder; chronic obstructive pulmonary disease;
 KW CNS disorder; diabetes; myocardial infarction; ischaemic heart disease;
 KW arrhythmia; hypertensive; Alzheimer's disease; Parkinson's disease;
 KW peripheral pain; chronic pain.
 XX OS Homo sapiens.
 XX WO2003000873-A2.
 XX PD 03-JAN-2003.
 XX 21-JUN-2002; 2002WO-EP006879.
 XX 25-JUN-2001; 2001US-0300071P.
 XX 16-NOV-2001; 2001US-0331447P.
 XX 07-DEC-2001; 2001US-0336693P.
 XX (FARB) BAYER AG.
 XX Xiao Y;
 XX WPI; 2003-201424/19.
 DR N-PSDB; AAL51590.
 XX New serine/threonine protein kinase NEK1 gene and protein, useful for
 PT identifying modulators of serine/threonine protein kinase NEK1 activity,
 PT and in gene therapy for treating cancer, diabetes, heart failure or
 PT Alzheimer's disease.
 XX Claim 1; Fig 2; 156pp; English.

XX The invention comprises the amino acid and coding sequence of the human
 CC serine/threonine protein kinase NEK1. The DNA and protein sequences of
 CC the invention are useful for modulating the activity of serine/threonine
 CC kinase NEK1 in a disease, such as: cancer (particularly colon cancer);
 CC cardiovascular disorders; central nervous system (CNS) disorders;
 CC diabetes; and chronic obstructive pulmonary disease. In particular the
 CC DNA and protein sequences of the invention are useful for treating:
 CC congestive heart failure; myocardial infarction; ischaemic heart disease;
 CC arrhythmia; hypertensive; Alzheimer's disease; Parkinson's disease; and
 CC peripheral or chronic pain. The present amino acid sequence represents
 CC the human serine/threonine protein kinase NEK1 of the invention
 XX Sequence 1242 AA;
 SQ
 Query Match 30.6%; Score 1105.5; DB 6; Length 1242;
 Best Local Similarity 33.7%; Pred. No. 2.4e-77;
 Matches 283; Conservative 121; Mismatches 208; Indels 229; Gaps 22;
 QY 1 MDKYDVIKAGGATGKAYLAKGSDSKHCVIKENFEKMPIQEKXKASKEVILLKMKH 60
 Db 1 MEKYVRLQKIGSGFGKALVASTEDGRQYVIKEINISMSKREESRREVAVLANMKH 60
 QY 61 PNIVAFNFSQNGRLFIWMEYCDGDLMKRINRGVLFSEDDQLGWFWQISLGLKHH 120
 Db 61 PNIVQRESFEENGSLIYMDYCEGGDLFKRINAGKGVLFQEDQLDWFVQICLAKHHV 120
 QY 121 DRKILHRTDKAONIFLSKNGMVAKLDFGIARVLNLSMELARTCIGTPYILSPETCONKP 180
 Db 121 DRKILHRTDKSONIFLTQDGV-QLGDFGIARVLNSTVELARTCIGTPYILSPETCONKP 179
 QY 181 YNKTDIWSLGCVLVELCTLKHPFEGNNLQQLVKICQAHFAPISPGFSRELHSLISQLF 240
 Db 180 YNKSIDIWALGCVLYELCTLKHAPEAGSMKNLVLIISGSPFVSHYSLYDLRLSVLSQLF 239
 QY 241 QVSPDRPSINSILKRPPLNLIPLYLTPVIOEFSHMLICRAGA---PASRHAG----293
 Db 240 KNPDRPSVNSILEKGFIAKRIEKLSPQIAEFCILTKFSKFGSPIPAKRPASGQNS 299
 QY 294 -----KWQCKIKQVRFQKCPSPRSISVIPKRNAI 325
 Db 300 ISVMPAQKITKPAKYGIPLAVKYGDKLEHKKPLQKHK-QAHTPEKRYNTGEERKI 358
 QY 326 -----LHRNE-----WRPAGAKQA 340
 Db 359 SEEAARKRRLEFIEKKQKQDIISLMKAEQMKEKQERLERINRARPQGNRVLSAGS 418
 QY 341 RSIKM-----IERPKIAAVCGHYDYVAOLDMLRRRAHKPSVHPHQENTGVEDYG--391
 Db 419 GEVKAPFLGSGGTIAPSPSSRGQYEHYHAIFDQMQQRAEDN-----EAKWKREIYGRG 473
 QY 392 ----QETHRGPSPSQWPAEYLQKFEAQYKLVKVEKQLGL-----RPSSAEPNPN 437
 Db 474 LPERQKQLAVERAKQVEEFLQKREAMQKARAEHGMILQNLAAVMYGGPSS-----527
 QY 438 QRQELRSNGEPRFOELPRKNEKQEWKQLEIRQYHNDKMEIRKQM-----488
 Db 528 -----SRGKPRNKE-----EVLARLQIRLQNFNERQQLAKLRGKKEANH 572
 QY 489 --GREPEENKISHKTYLVKNSLPVHQDASEGEAPVDIEKDLQKMLQNTKESKNEPQ 546
 Db 573 SEGQSGSEADMRK-----KIESLKAHANA---RAAV--LKEQLERKKEAYEREKVVWE 623
 QY 547 KYKAKKGVKF-----EINLQKISDENILOEE- 573
 Db 624 EHLVAKGVKSDVSPPLQGHETGGSPSQQMRSVISVTSALKKEGVDSLSLDTRETSEEM 683
 QY 574 -----EAMDIPNETLTFEDGMKFKE-----YECVKEHGYDT- 605
 Db 684 QKTNNAISSKREILRLNENLKAQDEKQKQNLSDTFEINVHEDAKHEKEKSVSDRKK 743
 QY 606 -KAFELKHP-----EAGFSTQTVAAVGN-----PROWDGGAPQTLQMAVAD 648

Db 744 WEAGQGLVPLDELTDTSFSTERTHTVGEVILKGPNGSPRAWKSPDTSVLKILGEAE 803
Qy 649 I 649
Db 804 L 804
RESULT 15
ID AAU07102 standard; protein; 1214 AA.
XX AC AAU07102;
XX DT 24-OCT-2001 (first entry)
XX DE Human novel human protein, NHP #2.
XX KW Human; novel human protein; NHP; breast cancer; prostate cancer;
XX KW immunogen; antibody; gene therapy; antisense.
XX OS Homo sapiens.
XX PN WO200161016-A2.
XX PD 23-AUG-2001.
XX PF 15-FEB-2001; 2001WO-US0053356.
XX PR 18-FEB-2000; 2000US-0183582P.
XX PR 22-FEB-2000; 2000US-0184014P.
XX PA (LEXI-) LEXICON GENETICS INC.
XX PI Walke DW, Hu Y, Nepomnichy B, Turner CA, Zambrowicz B;
XX WPI; 2001-502793/55.
XX DR N-PSDB; AAS11558.
XX PT Isolated nucleic acids encoding novel human proteins useful for the
XX treatment of disease and as probes for testing and detection.
XX PS Claim 4; Page 37-39; 69pp; English.
XX The invention relates to novel human proteins (NHP) and the nucleic acids
XX encoding them. The nucleic acids encode mammalian transporter proteins
XX and are useful for the treatment (e.g. by gene therapy or antisense
XX technology) of any of a wide variety of symptoms associated with
XX biological disorders (e.g. breast and prostate cancer) or imbalances and
XX as probes for the identification, selection and validation of novel
XX molecular targets for drug discovery. The proteins may be used to raise
XX anti-NHP antibodies. The present sequence represents an NHP of the
XX invention
XX SQ Sequence 1214 AA;

Query Match 29.9%; Score 1081.5; DB 4; Length 1214;
Best Local Similarity 34.4%; Pred. No. 1.8e-75;
Matches 274; Conservative 129; Mismatches 226; Indels 167; Gaps 21;
Qy 1 MKDYVIRKAGGAGKAVLAGKSDSKHCVIKENFEKMPIOEKEASKSVILLKMKH 60
Db 1 MEKYVRLQIGESFGKALVKSTEDGQYVIKEINIRMSGSKEREERREAVLANMKH 60
Qy 61 PNIVAFNFSQENGRLFIVMEYCDGDLMKRINRQGVLFSDQILGWVQISLGLKH 120
Db 61 PNIVQRESFEENGSIYVMDYCEGGDLFKRINACKGVLFQEDQLDFVQICLAKH 120
Qy 121 DRKILHRDIKAQNIFLSKNGVAKLGDGFIARVLNNSMELARTCIGTPYILSPICQ 180
Db 121 DRKILHRDIKSNIFLTKDGTV-QLGDFGIARVLNSTVELARTCIGTPYILSPIC 179
Qy 181 YNNKTDIWSLGVLYELCTLKHFAGSMKNVLKIIISGFPFVSLHYSYDLRLSLV 239
Qy 241 QVSPRORPSINILKRPFLLENLIPKYLTPRIVIOEESHMLICRAGA---PASRHAG--- 293
Db 240 KRNPDRPVSNSILEKGFIAKRIEFLSPOLIAEBFCLKTFSGSQTPAKKPASQONS 299
Qy 294 -----KVQCKI QKVRFOCKPPSRISVPIKRNAI 325
Db 300 ISVMPACKITPAKYGIPLAYKKYGDKKLHEKKPLQKHK-QAHQTPKRVNTGEERRKI 358
Qy 326 -----LHRNE-----WRPPAGAQA 340
Db 359 SEEAARKRLFEIEKKEKKQKQIIISLMKABQMKRQERLERINRAREQGNVLSAGGS 418
Qy 341 RSIXM-----IERPKIAVCGHYDYIAQLDMLRRRAHKPSYHPIQENTGVEDYG-- 391
Db 419 GEVKAPFLGSGGTIAPSSFSRGQYEHYHAI PQMOQQAEDN-----BAKWKEIYVGRG 473
Qy 392 -----QETRHGSPSPQWPAEYLQKFAQYKLVKVE-----KQLGLRPPSAEPNYNQ 438
Db 474 LPERQGLAVERAKQVEEFLOKREAMQNKARAEHGMVYLARLQIRLQ-----NFNE 527
Qy 439 RQELRS--NGEERPFQELPFRKNEMKEQYWKQLEBIRQOYHND---MKEIRKMGREPE 493
Db 528 RQIKAKLRGEKKEANHSEGESEBADMWRKKIESLKAHANARAVALKEQLERKKEAY 587
Qy 494 ENSKISHKTYLVKKS-----NLPVQDASEGEAPVQ-----DIEKDLKQM-----RLQN 537
Db 588 EREKVVWEHLVAKYKSDVSPFLQGHETGGSPSKQMRSVISVTSALKEVGVDSSLTD 647
Qy 538 TKESKNPEQK---YKAKGVKFEINLDKICISDENILQEBEAMDIPNETLTTFEDGMKFXE 593
Db 648 TRETSEEMQKTNNNAISSKREILRLN-----ENLKAQDEKGMQNLSDTFFINVHEDA 700
Qy 594 YECVKEHGDYTDK-----AFKELHCP-----EAGFSTQTVAAVGN-----RRQWD 633
Db 701 KEHEKEKSVSSDRKWEAGGQVLPLDELTDTSFSTERTHTVGEVILKGPNGSPRAWG 760
Qy 634 GGAPQTLLOMMAVADI 649
Db 761 KSPTDSVLKILGEAEL 776

Search completed: October 8, 2004, 17:07:50
Job time : 70 secs

Db 180 YNNKTDIWSLGVLYELCTLKHFAGSMKNVLKIIISGFPFVSLHYSYDLRLSLV 239
Qy 241 QVSPRORPSINILKRPFLLENLIPKYLTPRIVIOEESHMLICRAGA---PASRHAG--- 293
Db 240 KRNPDRPVSNSILEKGFIAKRIEFLSPOLIAEBFCLKTFSGSQTPAKKPASQONS 299
Qy 294 -----KVQCKI QKVRFOCKPPSRISVPIKRNAI 325
Db 300 ISVMPACKITPAKYGIPLAYKKYGDKKLHEKKPLQKHK-QAHQTPKRVNTGEERRKI 358
Qy 326 -----LHRNE-----WRPPAGAQA 340
Db 359 SEEAARKRLFEIEKKEKKQKQIIISLMKABQMKRQERLERINRAREQGNVLSAGGS 418
Qy 341 RSIXM-----IERPKIAVCGHYDYIAQLDMLRRRAHKPSYHPIQENTGVEDYG-- 391
Db 419 GEVKAPFLGSGGTIAPSSFSRGQYEHYHAI PQMOQQAEDN-----BAKWKEIYVGRG 473
Qy 392 -----QETRHGSPSPQWPAEYLQKFAQYKLVKVE-----KQLGLRPPSAEPNYNQ 438
Db 474 LPERQGLAVERAKQVEEFLOKREAMQNKARAEHGMVYLARLQIRLQ-----NFNE 527
Qy 439 RQELRS--NGEERPFQELPFRKNEMKEQYWKQLEBIRQOYHND---MKEIRKMGREPE 493
Db 528 RQIKAKLRGEKKEANHSEGESEBADMWRKKIESLKAHANARAVALKEQLERKKEAY 587
Qy 494 ENSKISHKTYLVKKS-----NLPVQDASEGEAPVQ-----DIEKDLKQM-----RLQN 537
Db 588 EREKVVWEHLVAKYKSDVSPFLQGHETGGSPSKQMRSVISVTSALKEVGVDSSLTD 647
Qy 538 TKESKNPEQK---YKAKGVKFEINLDKICISDENILQEBEAMDIPNETLTTFEDGMKFXE 593
Db 648 TRETSEEMQKTNNNAISSKREILRLN-----ENLKAQDEKGMQNLSDTFFINVHEDA 700
Qy 594 YECVKEHGDYTDK-----AFKELHCP-----EAGFSTQTVAAVGN-----RRQWD 633
Db 701 KEHEKEKSVSSDRKWEAGGQVLPLDELTDTSFSTERTHTVGEVILKGPNGSPRAWG 760
Qy 634 GGAPQTLLOMMAVADI 649
Db 761 KSPTDSVLKILGEAEL 776

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2004, 17:02:16 ; Search time 52 Seconds
(without alignments)
4144.210 Million cell updates/sec

Title: US-09-940-921B-2
Perfect score: 3616
Sequence: 1 MDKYDVIRKAGGAFKAYL.....SVSRQEGTKDPYSPVLILM 693

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL.25.*

- 1: sp_arChaea.*
- 2: sp_bacteria.*
- 3: sp_fungi.*
- 4: sp_human.*
- 5: sp_invertebrate.*
- 6: sp_mammal.*
- 7: sp_mhc.*
- 8: sp_orcanelle.*
- 9: sp_phage.*
- 10: sp_plant.*
- 11: sp_prodent.*
- 12: sp_virus.*
- 13: sp_vertebrate.*
- 14: sp_unclassified.*
- 15: sp_rvirus.*
- 16: sp_bacteriap.*
- 17: sp_archaeap.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query	Score	Match	Length	ID	Description
1	1778.5	49.2	614	11	Q8CCJ0	Q8ccj0 mus musculu
2	1778.5	49.2	627	11	Q7TSC3	Q7tsc3 mus musculu
3	1135.5	31.4	336	11	Q8C6N6	Q8c6n6 mus musculu
4	1006.5	27.8	424	11	Q8CD72	Q8cd72 mus musculu
5	986	26.7	302	11	Q8BSB6	Q8bsb6 mus musculu
6	799	22.1	489	4	Q8J023	Q8j023 homo sapien
7	797.5	22.1	509	11	Q99K72	Q99k72 mus musculu
8	695.5	19.2	849	13	Q7T299	Q7t299 brachydanio
9	675.5	18.7	1123	5	Q8G106	Q8g106 dictyosteli
10	617.5	17.1	555	5	Q9N9C3	Q9n9c3 leishmania
11	602	16.6	645	4	Q8NG66	Q8ng66 homo sapien
12	588	16.3	637	6	Q8WN08	Q8wn08 macaca fasc
13	584	16.2	482	4	Q8TBY1	Q8tby1 homo sapien
14	583	16.1	470	4	Q8NG65	Q8ng65 homo sapien
15	583	16.1	628	11	Q8C0Q4	Q8c0q4 mus musculu
16	575	15.9	579	5	Q95XQ3	Q95xq3 caenorhabdi

17	571	15.8	365	11	Q8BW62	Q8bw62 mus musculu
18	569	15.7	357	5	Q01775	Q01775 caenorhabdi
19	567.5	15.7	841	5	Q9VC32	Q9vc32 drosophila
20	566	15.7	438	11	Q8BLS6	Q8bls6 mus musculu
21	557	15.4	442	13	Q7ZVE3	Q7zve3 xenopus lae
22	554	15.3	442	13	Q9W622	Q9w622 xenopus lae
23	550.5	15.2	568	10	Q8XA66	Q8xa66 arabidopsis
24	546.5	15.1	389	13	Q9W623	Q9w623 xenopus lae
25	546	15.1	366	11	Q8BP64	Q8bp64 mus musculu
26	542	15.0	561	5	Q76134	Q76134 tetrahymena
27	540.5	14.9	443	11	Q91Z18	Q91z18 mus musculu
28	539.5	14.9	491	4	Q86XH2	Q86xh2 homo sapien
29	534.5	14.8	326	4	Q7Z634	Q7z634 homo sapien
30	534.5	14.8	384	4	Q96QN9	Q96qn9 homo sapien
31	534.5	14.8	943	10	Q94C05	Q94cu5 oryza sativ
32	534	14.8	440	13	Q7ZUN2	Q7zun2 brachydanio
33	532	14.7	609	10	Q947T1	Q947t1 lycopersico
34	531.5	14.7	443	11	Q921N9	Q921n9 mus musculu
35	529	14.6	549	5	Q819T8	Q819t8 tetrahymena
36	528.5	14.6	416	10	Q9LT35	Q9lt35 arabidopsis
37	525	14.5	618	10	Q8S3U7	Q8s3u7 chlamydomon
38	523	14.5	555	10	Q8RXT4	Q8rxt4 arabidopsis
39	518.5	14.3	621	10	Q8SA64	Q8sa64 populus x c
40	511.5	14.1	393	11	Q91XQ1	Q91xq1 rattus norv
41	507.5	14.0	606	10	Q9CAU7	Q9cau7 arabidopsis
42	495	13.7	1057	5	Q81SD5	Q81sd5 plasmodium
43	483	13.4	356	5	Q8N3P1	Q8n0p1 paramecium
44	476	13.2	735	5	Q9W3N8	Q9w3n8 drosophila
45	475.5	13.1	299	5	Q8T755	Q8t755 branchiosto

ALIGNMENTS

RESULT 1
Q8CCJ0
ID Q8CCJ0 PRELIMINARY; PRT; 614 AA.
AC Q8CCJ0;
DT 01-MAR-2003 (TREMBLrel. 23, Created)
DT 01-MAR-2003 (TREMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TREMBLrel. 25, Last annotation update)
DE Weakly similar to protein kinase nek1.
GN BBO49667.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Cerebellum;
RX MEDLINE=22354683; PubMed=12466851;
RA The FANTOM Consortium,
RT "Analysis of the mouse transcriptome based on functional annotation of
RL Nature 420:563-573(2002).";
DR EMBL; AK032672; BAC27980.1; --
DR MGD; MGI:12142824; BB049667.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; pkinase; I.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR SMART; SM00219; TYKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.

DR PROSITE; P300108; PROTEIN KINASE ST; 1.
SQ SEQUENCE 614 AA; 70032 MW; DA95120EA45D71BC CRC64;

Query Match 49.2%; Score 1778.5; DB 11; Length 614;
Best Local Similarity 54.1%; Pred. No. 5.7e-107;
Matches 366; Conservative 55; Mismatches 79; Indels 177; Gaps 8;

QY 1 MDKYDVIRKAGCGAFGKAYLAKGSDSKHCVIKEINFEKMPITQEKESKKEVILEKMKH 60
DB 1 MDNFHLIKIIGSTGKGVYLAQKXSESHCVIKEISLTG-----EKEASKNEVILLARMEH 56

QY 61 PNIVAFNSFQNGRLFIWMEYCDGDLMLKRNORGVLFSDQILGFWQISLGLKHH 120
DB 57 PNIVTFSSFQNGRLFIWMEYCDGDLMLKRNORGVLFSDQILGFWQISLGLKHH 116

QY 121 DRKILHRDIKAQNIIFLSKNGMVAKLDFGIARVLNNSMELARTCTGTYYLSPEICQNK 180
DB 117 DRKILHRDIKSNQIFLSKNGMVAKLDFGTARTLNSMELAQTCAGTYYLSPEICQNR 176

QY 181 YNKTDIWSLGVLYELCTLKHPPFGNNLQQLVLCQAHFAPISPGFSRELHSLISOLF 240
DB 177 YNKTDIWSLGVLYELCTLKHPPFSNNFHLVLKICQGRVAPISPHFSRDLQSLIPOLF 236

QY 241 QVSPDRPSINSILKRPFLNIPKYLTPETVQEBFBSHMLICRAGAPASRHAGKVQCK 300
DB 237 RVSPQDRPSVTSLLKRPFLNIPKYLTPETVQEBFBSHMLICRAGAPASRHAGKVQCK 267

QY 301 IQKVFQKCPKPSRISVPIKENAILHRNEWPPAGAQAQARSIMIERPKIAAVCGHYD 360
DB 268 -----CSR----- 271

QY 361 YYAQLDMLRRRAHKFSYHPIQENTGVEDYGOETRHGFS-----SOWPAEYLQKFEAQ 416
DB 272 -----IQSHAH-----MENNAI-----GPTACWRVSPWSAAVLRKFEAQ 307

QY 417 YKLKVEKQILGRPSAEPNYQRLNGEBEPFQELPFRKNEMKQEYKQLEIEIQ 476
DB 308 YKLKVERQLGRPSSVEPHNPEGEKQLQSHWEETKQELQYRKNKQKQDEYKQLEIEIQ 367

QY 477 YHNDKKEIRKQGRPEPNSKISHKTYLVKKNLPHQDASEGEAPVODIEKDLQKMLQ 536
DB 368 YHNDKKEIRKQGRPE----- 382

QY 537 NTKESKNPEQKYAKKGVKVFENLQKICSDENILOEEAMDIPTNLTFFDGMKEVEY 596
DB 383 -----LKRUVKFEISLQKICSEEDTVQENEAVDKLNATLSFEDGTKEQHR 429

QY 597 VKEHGYTDKAFKELHCPAGESTQTVAAVGNRRQWDGAPQTLQMAVADITSTCTPTG 656

RESULT 2
Q7TSC3 PRELIMINARY; PRT; 627 AA.

AC Q7TSC3
DT 01-OCT-2003 (TRENBLrel. 25, Created)
DT 01-OCT-2003 (TRENBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Hypothetical protein.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
Diatchenko L., Marusina K., Farmer A.A., Rubin G.W., Hong L.,
Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E.,
Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullahy S.J.,
Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
Krzywinski M.I., Skalska U., Smailus D.E., Schnerch A., Schein J.E.,
Jones S.J., Marra M.A.;
"Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).

RP SEQUENCE FROM N.A.
RC TISSUE=Olfactory epithelium;
RL Submitted (JUN-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL: BC053516; AAH53516.1; -
DR Hypothetical protein.
KW Hypothetical protein.
SQ SEQUENCE 627 AA; 71790 MW; AD5789B6621ACA2P CRC64;

Query Match 49.2%; Score 1778.5; DB 11; Length 627;
Best Local Similarity 54.1%; Pred. No. 5.8e-107;
Matches 366; Conservative 55; Mismatches 79; Indels 177; Gaps 8;

QY 1 MDKYDVIRKAGCGAFGKAYLAKGSDSKHCVIKEINFEKMPITQEKESKKEVILEKMKH 60
DB 1 MDNFHLIKIIGSTGKGVYLAQKXSESHCVIKEISLTG-----EKEASKNEVILLARMEH 56

QY 61 PNIVAFNSFQNGRLFIWMEYCDGDLMLKRNORGVLFSDQILGFWQISLGLKHH 120
DB 57 PNIVTFSSFQNGRLFIWMEYCDGDLMLKRNORGVLFSDQILGFWQISLGLKHH 116

QY 121 DRKILHRDIKAQNIIFLSKNGMVAKLDFGIARVLNNSMELARTCTGTYYLSPEICQNK 180
DB 117 DRKILHRDIKSNQIFLSKNGMVAKLDFGTARTLNSMELAQTCAGTYYLSPEICQNR 176

QY 181 YNKTDIWSLGVLYELCTLKHPPFGNNLQQLVLCQAHFAPISPGFSRELHSLISOLF 240
DB 177 YNKTDIWSLGVLYELCTLKHPPFSNNFHLVLKICQGRVAPISPHFSRDLQSLIPOLF 236

QY 241 QVSPDRPSINSILKRPFLNIPKYLTPETVQEBFBSHMLICRAGAPASRHAGKVQCK 300
DB 237 RVSPQDRPSVTSLLKRPFLNIPKYLTPETVQEBFBSHMLICRAGAPASRHAGKVQCK 267

QY 301 IQKVFQKCPKPSRISVPIKENAILHRNEWPPAGAQAQARSIMIERPKIAAVCGHYD 360
DB 268 -----CSR----- 271

QY 361 YYAQLDMLRRRAHKFSYHPIQENTGVEDYGOETRHGFS-----SOWPAEYLQKFEAQ 416
DB 272 -----IQSHAH-----MENNAI-----GPTACWRVSPWSAAVLRKFEAQ 307

QY 417 YKLKVEKQILGRPSAEPNYQRLNGEBEPFQELPFRKNEMKQEYKQLEIEIQ 476
DB 308 YKLKVERQLGRPSSVEPHNPEGEKQLQSHWEETKQELQYRKNKQKQDEYKQLEIEIQ 367

QY 477 YHNDKKEIRKQGRPEPNSKISHKTYLVKKNLPHQDASEGEAPVODIEKDLQKMLQ 536
DB 368 YHNDKKEIRKQGRPE----- 382

QY 537 NTKESKNPEQKYAKKGVKVFENLQKICSDENILOEEAMDIPTNLTFFDGMKEVEY 596
DB 383 -----LKRUVKFEISLQKICSEEDTVQENEAVDKLNATLSFEDGTKEQHR 429

QY 597 VKEHGYTDKAFKELHCPAGESTQTVAAVGNRRQWDGAPQTLQMAVADITSTCTPTG 656
DB 430 KEEHEDYTDRAPEELCGPEAEFFQVDAENRRQWDGAPHTLLRIMAMADVSTCTPTM 489

QY 657 PDSVSLV-----SRQEK 671
DB 490 PDGGQIVNVEGSENGK 506

QY	657	PDSSEVLSV--SRQCK 671		:	:	:		
Db	490	PDGQVIVWEGSVENGK 506		:	:	:		
RESULT 3								
Q8C6N6		PRELIMINARY; PRT; 336 AA.						
ID	Q8C6N6							
AC	Q8C6N6;							
DT	01-MAR-2003 (T-EMBLrel. 23, Created)							
DT	01-MAR-2003 (T-EMBLrel. 23, Last sequence update)							
DT	01-OCT-2003 (T-EMBLrel. 25, Last annotation update)							
DE	Weakly similar to protein kinase nekl.							
DB	BB049667.							
OS	Mus musculus (Mouse).							
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;							
OC	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.							
OX	NCBI_TaxID=10090;							
XP	[2]							
RN	SEQUENCE FROM N.A.							
RP	RC STRAIN=C57BL/6J; TISSUE=Oviduct;							
RC	MEDLINE=22354683; PubMed=12466851;							
RX	RA The FANTOM Consortium.							
RA	the RIKEN Genome Exploration Research Group Phase I & II Team;							
RT	"Analysis of the mouse transcriptome based on functional annotation of							
RT	60,770 full-length cDNAs.";							
RL	Nature 420:563-573 (2002).							
DR	EMBL; MG1:2142824; BAC35677.1; .							
DR	MGM; MG1:2142824; BB049667.							
DR	GO; GO:000524; F:ATP binding; IEA.							
DR	GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.							
DR	GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.							
DR	GO; GO:0006468; F:protein amino acid phosphorylation; IEA.							
DR	InterPro; IPR000713; Prot_kinase.							
DR	InterPro; IPR002290; Ser_thr_kinase.							
DR	InterPro; IPR008271; Ser_thr_kin AS.							
DR	InterPro; IPR001245; Tyr_kinase.							
DR	Pfam; PF00069; pkinase; 1.							
DR	PRINTS; PR00109; TYRKINASE.							
DR	ProDom; PD000001; Prot_kinase; 1.							
DR	SMART; SM00220; S_TKc_1.							
DR	SMART; SM00219; TyrKc; 1.							
DR	PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.							
DR	PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.							
DR	PROSITE; PS00108; PROTEIN_KINASE_ST; 1.							
DR	SEQUENCE 336 AA; 38757 NW; 6BDD37D72E0143B CRC64;							
QY	Query Match	31.4%;	Score	1135.5;	DB	11;	Length	336;
Db	Best Local Similarity	77.0%;	Pred. No.	1.2e-65;				
Db	Matches	217;	Conservative	22;	Mismatches	36;	Indels	7;
Db								
QY	1	MDKYDVKAIGQAFGKAYLAKGSDSKHCYKINFERMPIQEKASKKEVILLERKXG 60						
Db	1	MDNFHLIKIIGETGKVLAKDSESSHCYKISLTK---EKEASKNEVILLARMEH 56						
QY	61	PNIVAFNSFQNGRLFIWMEYCGDGLMKENRQGVLSDDOILCNFVQISLGKXTH 120						
Db	57	PNIVTFSSQNGRLFIWMEYCDGDDLMQRIQRQGMFSEDOILCNFVQISLGKXTH 116						
QY	121	DRKILHRDIAKQNIIFLSKNGWAKLGDGFIARVLNNSMELARTICGTFPYLSPEICQNP 180						
Db	117	DRKILHRDIKQNIIFLSKNGWAKLGDGFTARTLNDSMELAQTCAGTFYLSPEICQNP 176						
QY	181	YNNKTDIWSLGCVLVELCTLKHPPEGNNLQVLKICQAHFAPISPGFSRELHSLISOLF 240						
Db	177	YNNKTDIWSLGCVLVELCTLKHPPESNNFHLVLIQGRVAPISPHFSRDLQSLIPQLF 236						
QY	241	QVSPDRPSINSILKRPFLNLIKPLTPEVIQ---EEFGHM 279						
Db	237	RVSPDRPSVTLKKRPFLNLIKPLTARSLYPEVCSRRIOASHAM 278						

```

RESULT 4
Q8CD72      PRELIMINARY;      PRT;      424 AA.
ID          Q8CD72
AC          Q8CD72;
DT          01-MAR-2003 (TrEMBLrel. 23, Created)
DT          01-MAR-2003 (TrEMBLrel. 23, Last sequence update)
DT          01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE          NIWA.
GN          NEKL.
OS          Mus musculus (Mouse).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX          NCBI_TaxID=10090;
OX          [1]
RN          SEQUENCE FROM N.A.
RP          STRAIN=C57BL/6J; TISSUE=Testis;
RX          MEDLINE=22354683; PubMed=12468851;
RA          the FANTOM Consortium,
RA          the RIKEN Genome Exploration Research Group Phase I & II Team;
RT          "Analysis of the mouse transcriptome based on functional annotation of
RT          60,770 full-length cDNAs.";
RL          Nature 420:563-573(2002).
DR          EMBL; AK031330; BAC27350.1; -.
DR          MGD; MGI-97303; Nekl.
DR          GO; GO:0005524; F:ATP binding; IEA.
DR          GO; GO:004674; F:protein serine/threonine kinase activity; IEA.
DR          GO; GO:004713; F:protein-tyrosine kinase activity; IEA.
DR          GO; GO:0006458; P:protein amino acid phosphorylation; IEA.
DR          InterPro; IPR000719; Prot kinase.
DR          InterPro; IPR002290; Ser_thr_kinase.
DR          InterPro; IPR008271; Ser_thr_pkin_AS.
DR          InterPro; IPR001245; Tyr_kinase.
DR          Pfam; PF00069; Pkinase; 1.
DR          PRINTS; PRC0109; TIRKINASE.
DR          PRODOM; PD000001; Prot_kinase; 1.
DR          SMART; SM00220; S_TKC; 1.
DR          SMART; SM00219; TYRK; 1.
DR          PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR          PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR          PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR          SEQUENCE      424 AA; 46635 MW; 416AD0E55C13CEFE CRC64;

Query Match      27.8%; Score 1006.5; DB 11; Length 424;
Best local Similarity 52.2%; Pred. No. 3.4e-57;
Matches 198; Conservative 64; Mismatches 84; Indels 33; Gaps 5;

Qy          1 MDKYDVILKALGQAGFKAYLAKGSDSKHCVKEINPEKPIQKESKKEVILEKMKH 60
Db          1 MEKVRLQKIGESFGKAVLVKSTEDGRHVYVKEINISRMSDKERQSRREVAVLANKH 60

Qy          61 PNIVAFNFSQENGRLFIVMEYCDGDLKRIINRQGVLFSEQILQWFOVQISGLKHII 120
Db          61 PNIVQYKSEFEENGSLYIVMDYCEGGDLFKRIINAQKALPQEDQLDFWQVQICLALKHVH 120

Qy          121 DRKILHRDIAQNIFFLSKNGVAKLGDGFIARVLNNSMEIARTCIGPYVLSPEICNKP 180
Db          121 DRKILHRDIAKSNIFLTQGVVQLGDFGIARVLNSTVELARTCIGHPYVLSPEICNKP 179

Qy          181 YNNKTDITWSLGCVLYELCTLKHFFEGNNLQOLVLKICQAHAFIPSPGFSRELHSLISQLF 240
Db          180 YNNKSDITWALGCVLYELCTLKHAFEGNKNVLVLIKISGSPFPVSPHYSDLRSLLSQLF 239

Qy          241 QVSPDRPRSINSILKRPFLLENLPKYLTPVEIOEFSHMLICRAG----- 285
Db          240 KRPVDRPSPVNSILEKGFIAKRIEKLSPOLIAEEFCLKTLSEKGPQLPKRPSAQGV 299

Qy          286 -----APASRHAGKVQVCKIKVRFQCKPPRPRIS--VPKRNAILHRNWRP 333
Db          300 SSFVPAGKIKTPAAKYGVPLTYKYGDKLLKLEKPPPKHQAHQIPVK---MNSGERK 356

Qy          334 PAG--AOKARSIKWIRPK 350

```

```

Db 357 KMSEBAKKRLEFIEKEK 375
RESULT 5
QBSB6
ID QBSB6 PRELIMINARY; PRT; 302 AA.
AC QBSB6;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE NIMA.
GN NEK1.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=C57BL/6J; TISSUE=Embryo;
RX MEDLINE=22354683; PubMed=12466851;
RA The PANTOM Consortium,
RA "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573 (2002).
DR EMBL; AK034754; BAC28822.1; -.
DR MGD; MGI:97303; Nek1.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
SQ SEQUENCE 302 AA; 34328 MW; 6D5B14D7809CEB341 CRC64;

Query Match 26.7%; Score 966; DB 11; Length 302;
Best Local Similarity 62.4%; Pred. No. 9.5e-55;
Matches 181; Conservative 47; Mismatches 60; Indels 2; Gaps 2;

QY 1 MDKYVIRKAGGAFKAYLAGKSDSGCHYKKEINFERMPIQKEASKEVILLKMKH 60
Db 1 MEKYVRLQIGSGFGKAVLVKSTEDGRHYVYKEINISRMDSKQERQSRREAVLANKH 60
QY 61 PNIVAFNSFQNGRLFTVMEYCDGDLMKRINRGVLFSEDTLGLWFWVOISLGLKH 120
Db 61 PNIVQKSFENGSLYVMDYCEGDLFKRINAGKGLFQEDQLDWFVQICLAKKHV 120
QY 121 DRKILHRDIKAQNIFLSKNGMWAKLGDGFIARVLNNSMELARTCIGTPYLSPEICQNK 180
Db 121 DRKILHRDIKSQNIPLTKDGTV-QLGDFGIARVLNSTVELARTCIGTPYLSPEICENKP 179
QY 181 YNNKTDIWSLGCVLVELCTLKHPPGNNLQOLVKIQAHFAPISPGFSRELHLSIQLF 240
Db 180 YNNKSDIWSLGCVLVELCTLKHAFAGKWKLVLIISGSPFPSPHYSDRLSLSQLF 239
QY 241 QVSPDRPSINSILKRPFLNIPKYLTPVEIQEFSHMLICRACAPASR 290
Db 240 KRNPRDRPSVNSILEKGIKRIEFLSPQVC-VTLPLTLNCKVLVPKTR 288

RESULT 6
QBS023
ID QBS023 PRELIMINARY; PRT; 489 AA.
AC QBS023;
DT 01-MAR-2003 (T-EMBLrel. 23, Created)
DT 01-MAR-2003 (T-EMBLrel. 23, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE NIMA-related protein kinase 3.
GN NEK3.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Kimura M., Okano Y.;
RT "Molecular cloning of human NEK3.";
RL Submitted (OCT-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AS072828; BAC15599.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0004688; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR002290; Ser Thr kinase.
DR InterPro; IPR008271; Ser_thr_kin_AS.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase.
SQ SEQUENCE 489 AA; 55983 MW; FAB5DAAC01EFA88A CRC64;

Query Match 22.1%; Score 799; DB 4; Length 499;
Best Local Similarity 54.0%; Pred. No. 1.1e-43;
Matches 148; Conservative 51; Mismatches 73; Indels 2; Gaps 2;

QY 1 MDKYVIRKAGGAFKAYLAGKSDSGCHYKKEINFERMPIQKEASKEVILLKMKH 60
Db 1 MDKYVIRKAGGAFKAYLAGKSDSGCHYKKEINFERMPIQKEASKEVILLKMKH 60
QY 61 PNIVAFNSFQNGRLFTVMEYCDGDLMKRINRGVLFSEDTLGLWFWVOISLGLKH 120
Db 60 PNIVAFNSFQNGRLFTVMEYCDGDLMKRINRGVLFSEDTLGLWFWVOISLGLKH 119
QY 121 DRKILHRDIKAQNIFLSKNGMWAKLGDGFIARVLNNSMELARTCIGTPYLSPEICQNK 180
Db 120 KKEVLRDIKSNIPLTKQNGV-KLGDGFSARLLSNPMAFACTYVGTVPYVPEIWENLP 178
QY 181 YNNKTDIWSLGCVLVELCTLKHPPGNNLQOLVKIQAHFAPISPGFSRELHLSIQLF 240
Db 179 YNNKSDIWSLGCVLVELCTLKHPPGNNLQOLVKIQAHFAPISPGFSRELHLSIQLF 238
QY 241 QVSPDRPSINSILKRPFLNIPKYLTPVEIQE 274
Db 239 KRNPRDRPSATLLSRGIVARLVQKRLPEIIME 272

RESULT 7
QBSK72
ID QBSK72 PRELIMINARY; PRT; 509 AA.
AC QBSK72;
DT 01-JUN-2001 (T-EMBLrel. 17, Created)
DT 01-JUN-2001 (T-EMBLrel. 17, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Similar to NIMA (never in mitosis gene a)-related expressed kinase 3.
GN NEK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 OX NCBI_TaxID=10090;
 RN [1]_SEQUENCE FROM N.A.
 RA Strausberg R.;
 RL Submitted (MAR-2001) to the EMBL/GenBank/DBJ databases.
 CC -!- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
 DR EMBL; BC005411; AAH05411.1; -.
 DR HSSP; O63450; 1A06
 DR MGD; MG1:1344371; Nek3.
 DR GO; GO:0005274; F:ATP binding; IEA.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
 DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
 DR GO; GO:0016740; F:transferase activity; IEA.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc_1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 KW ATP-binding; Kinase; Serine/threonine-protein kinase; Transferase.
 SQ SEQUENCE 509 AA; 57109 MW; 333585D0588868A1 CRC64;

Query Match 22.1%; Score 797.5; DB 11; Length 509;
 Best Local Similarity 32.7%; Pred. No. 1.5e-43;
 Matches 201; Conservative 100; Mismatches 187; Indels 127; Gaps 17;

QY 1 MDKYDVIRKAIQAGAFKAYLAKGSDSKHCIVKEINFEKMPIQEAKSKVILLKMKH 60
 DB 1 MDNYTVLRVIGGSGRALLVLQESSNQTFAKKEIRLLK---SDQTSEKAVLLAKMKH 57

QY 61 PNIVAFNSFQENGLFTVMEYCDGDLMKRINRGVLFSEDTQLGFWFVQISLGLKHI 120
 DB 58 PNIVAFKESFEAGLYIVMEYCDGDLMKRIKQKGLFPEDTILNFIQICLVNHIH 117

QY 121 DRKILHRDIKAONIFLSKNGVMKALGDFGIARLVNLSMELARTCIGTPYVLSPEICQNK 180
 DB 118 KERVLRHLRIKSNVLETHNGKV-KLGDGSGARLSSPNMAFCTVGTPTVPPPIWENLP 176

QY 181 YNNKTDIWSLGLVVELCTLKHPPFGNNLQOLVLKICQAHFAP:SPGFSRELHLSIQLF 240
 DB 177 YNNKSDIWSLGLVVELCALKHPFGANSKXNLILKICQGP:HPALYSCKLQDLVKQML 236

QY 241 QVSPDRDPSINSILKRPLEMLIPKYLTPVIOEFSHMLICRAGAPASRAGKVQCK 300
 DB 237 KNPSPHRFSVTLLCRGSLAPLVKCLPPIRE-----YGEQILDEIK 280

QY 301 IQKVFQGGKPPRSRISVPIKENAILHNEWRPAGAKARSIKMIERPKIAAVCGHYDY 360
 DB 281 ISTDKNMK-----QDSNRVGEALGEANSASQOEER-GRKCSHTEL 321

QY 361 YQAQLDMLRRRAKESYHPIQENTGV-----EDVQETRHGSPSSQWPAEVLQKRFRAQ 416
 DB 322 -----ESTGTTT PGNALGRAARGNPGN---PQEGHRTSPASP 356

QY 417 YKLKVEKQLGLRPSAEPNQRQRLNGRNGEPRFQELPPKRNEMKEQYWKQLEELRQ 476
 DB 357 HRPVRERH---GPSNVEALEKASILTSS-----FTAEDRGSGSVIKYENARRQ 403

QY 477 YNDMMKETRKMGRPEPENSISKHTYLVKSNL-----PVHQDASEGAPVDIEKDL 530
 DB 404 WYREPEALLSMLKDADLSQ-AFQTYTYIRPGAEGFLKGLSELDTS-----DSVGDGL 456

QY 531 KOMRLQNTKESNPEQKYKAKGVKFEINLQK:SDENILQEEBAMDPINTELLTFEDGMK 590
 DB 457 DSVML-----DPE-----RFEPLDSE-EDTDFEEDNPDWLSE----- 489

QY 591 FKEYECVKEHGHDYTD 605
 DB 490 -----LKGVGYGD 498

RESULT 8
 QY71299 PRELIMINARY; PRT; 849 AA.
 ID QY71299
 AC QY71299;
 DT 01-OCT-2003 (TrEMBLrel. 25, Created)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Hypothetical protein.
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.
 OX NCBI_TaxID=7955;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RX MEDLINE=22386257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.P., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Frange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shcherchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
 RA Krzywinski M.I., Skalska U., Smallus D.E., Schnerch A., Schein J.E.,
 RA Jones S.J., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 PL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Kidney;
 RA Strausberg R.;
 RL Submitted (JUL-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC054633; AAH54633.1; -.
 KW Hypothetical protein.
 SQ SEQUENCE 849 AA; 95966 MW; E4E8B117A0147D7B CRC64;

Query Match 19.2%; Score 695.5; DB 13; Length 849;
 Best Local Similarity 27.2%; Pred. No. 1.1e-36;
 Matches 206; Conservative 131; Mismatches 280; Indels 139; Gaps 23;

QY 1 MDKYDVIRKAIQAGAFKAYLAKGSDSKHCIVKEINFEKMPIQEAKSKVILLKMKH 60
 DB 1 MDGYLFRVVVGXSGYGEVNLVRHKSQRKYQVTKKLNLRSSRRERRAAQEAQLSLKH 60

QY 61 PNIVAFNSFQ-ENGLFTVMEYCDGDLMKRINRGVLFSEDTQLGFWFVQISLGLKHI 119
 DB 61 PNIVMYRESWEGDCOLYIVMGCEGDLVHLKQKQKGLLPQVQVVEFVQIAMALQYL 120

QY 120 HDRKILHRDIKAONIFLSKNGVMKALGDFGIARLVNLSMELARTCIGTPYVLSPEICQNK 179
 DB 121 HEXHILHRDLKTONIFLTNTII-KVGDGLGIARLVENQNDMASTLIGTPTVYSPSPFSNK 179

QY 180 PYNKTDIWSLGLVVELCTLKHPPFGNNLQOLVLKICQAHFAP:SPGFSRELHLSIQSL 239
 DB 180 PYNKSDIWSLGLVVELCTLKHPPFGNNLQOLVLKICQAHFAP:SPGFSRELHLSIQSL 239

QY 240 FQVSPDRDPSINSILKRPLEMLIPKYLTPVIOE-----FSHMLICRAGAPASRHA 292

```
Db 240 LCKPEDRPDVKHILRQYIKHQISMFL--BATKEKTAKRKXKNAAGKLNAGSAGDASTKP 297
QY 293 GKVVO-KCKIKQVRQGGKPPRSRISVPIKNAIILHRENWRPPAGAKVARSIKMIERPXI 351
Db 298 NVQVQPQCLNSGSKTCG-----KKAEEIYLNRQKPCNGAWENVAPKHHMPKPS 345
QY 352 AAVCGHYDYAAQLDMLRRRAHKPSYHPQP-ENTGVEDVYQETRHGSPSPS-----QWPA 405
Db 346 PT-----RDIHNTGGQSIATISIDIEIQKQKAKPKPKSSHQNNLPS 389
QY 406 EYLQKFE-----AQOYKLGK-----VEKQGLRPPSAEPNINQORQLRNG----- 446
Db 390 VSKREKEDPGAPQTHPHKQYGVGGTDEKMSANASSITPKPADRTKPNKSKALDVSLD 449
QY 447 -----EPRPQELPFRKEM-----KEQYVWQLBEIROQYNDKKEIKWGRPEP 494
Db 450 IKDDTMKLLQAGVQDLPEYPTLHNAEQKTESQRTADNOSNAGVLKDAPTGSSTESE 509
QY 495 NSKISHKTYLVKSNLPVHQDASEGEAPVQD-----IEKDLKQVRLQNTKESKNPEQ 546
Db 510 GSLDSTEKLLPVPVIFNEPSESPPAALSDEGRPSSESPMSRQRRQKDKXSODLQ 569
QY 547 KYAKKGVKFEIN--LDKCIDENILQEE-----AMDIPNETLTFEDGMKFE 593
Db 570 -----NGIQEKVSRPLPLPEDGNSAQKDQSTPAGLSSTNKSTEQNRPLSARERRLKQ 624
QY 594 YE-----CVKEHG-DYTDKAFELH-----CPEAGPST--QTVAAGNRRQWDGGAPO 638
Db 625 SGNSTQQAASVAKRPSQVSLDSEQIHQDTPRAATTVDITTKERSKQSD----- 679
QY 639 TLLQWMAVADITSTCTGPDSESVLSVRQBGKTKD 674
Db 680 -----EDEC-----SSSTSTSERSEGDYRE 699

RESULT 9
Q86106
ID Q86106 PRELIMINARY; PRT; 1123 AA.
AC Q86106;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Similar to Mus musculus (Mouse). 13 days embryo male testis cDNA,
DE RIKEN full-length enriched library, clone:6030407f11 product:NIMA
DE (Never in mitosis gene a)-related expressed kinase 1, full insert
DE sequence.
OS Dictyostelium discoideum (slime mold).
OC Eukaryota; Mycetozoa; Dictyosteliida; Dictyostelium.
CX NCBI_TaxID=44689;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=22092622; PubMed=12097910;
RA Gloeckner G., Eichinger L., Szafranski K., Pachebat J., Dear P.,
RA Lehmann R., Baumann C., Parra G., April J.F., Guigo R., Kumpf K.,
RA Tunggal B., Cox E., Quail M.A., Platzer M., Rosenthal A., Noegel A.A.;
RT "Sequence and analysis of chromosome 2 of Dictyostelium discoideum.";
RL Nature 418:79-85(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=AX4;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR ENBL; AL358652; CAB94013.1; --
DR HSP; Q00534; IBI8.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002965; P-rich exten.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR008271; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_pkinase.
```

```
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PRO1217; PRICHEXTENSN.
DR PRINTS; PRO109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Kinase.
SQ SEQUENCE 1123 AA; 120808 MW; 69DFDESE2B0A3283 CRC64;

Query Match 18.7%; Score 675.5; DB 5; Length 1123;
Best Local Similarity 50.7%; Pred. No. 3.1e-35;
Matches 139; Conservative 46; Mismatches 84; Indels 5; Gaps 4;

QY 1 MDKYDVKAIGGAGATGKAYLAGKSDSKHCVIKENFEMPIQOEKAEKKEVILLKMKH 60
Db 1 MDKYBEIKTIGKSGFRALVKKRSDGLLLVKEINVMEMQPKERSDANNEVLLSLMDH 60
QY 61 PNIVAFFNSFQNGRLFTVMEYCDGDLMKRINRQ--RGVLFSDEQILGWVQISLGLKH 118
Db 61 ENHIGYDSFINGCLYIMEYANAGDINLEIKKRTLQNKTFSEFEILLSQICKALQY 120
QY 119 IHRKILHRDIKAQNFUS-KNG-MVAKLGDGFIARVLNNSMELARTCIGTPYILSPIC 176
Db 121 ISSRNILHRDLKTONIFLSIVNGDYFIKLGDFGIKILNSETSLASTVLGTGYILSPELI 180
QY 177 QN-KPVNKTDLWSLGCVLVELCTLKHPEGNLQOLVKICQAHFAPISPGFSRELHSL 235
Db 181 QNEKGYDHKSDIWSLGCVLVELTTLKHAFNAANLPAVLKILKGTYPPIPSHYSNDLNL 240
QY 236 ISQLFQVSPDRPSINSILKRPFLENLIPKYLTP 269
Db 241 ISSMLQIDPKNPSVNDILELFPINQYLGILPKP 274

RESULT 10
Q9N9C3
ID Q9N9C3 PRELIMINARY; PRT; 555 AA.
AC Q9N9C3;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE NEK-related serine/threonine-protein kinase nek1.
DE Pi408.02.
OS Leishmania major.
OC Eukaryota; Euglenozoa; Kinetoplastida; Trypanosomatidae; Leishmania.
CX NCBI_TaxID=5664;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RA Bothe G., Pohl T., Ivens A.C., Quail M., Rajandream M.A.,
RA Barrell B.G.;
RL Submitted (DEC-2000) to the EMBL/GenBank/DBSJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=Friedlin;
RX MEDLINE=98146435; PubMed=9477341;
RA Ivens A.C., Lewis S.M., Bagherzadeh A., Zhang L., Chan H.M.,
RA Smith D.F.;
RT "A physical map of the Leishmania major Friedlin genome.";
RL Genome Res. 8:135-145(1998).
CC -1- SIMILARITY: BELONGS TO THE SER/THR FAMILY OF PROTEIN KINASES.
DR ENBL; AL358652; CAB94013.1; --
DR HSP; Q00534; IBI8.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0006468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
```



```

Db 362 KIVEEK-----YENSRMELRSRNFQQLSDVHLHKTLLKG----- 399
QY 405 AYLOKFAQYQYKLVKVEKQLGRFPSSAEPNYNQQLASNGEPRFQELPFRK----- 458
Db 400 ---MEBKEQPEGRILSCSPQ-----DEDERWQGRE---ESDEPTLENLPSPQIPSM 448
QY 459 -----NEMKEQEQYWK 468
Db 449 LHELESIVEDATSDGLYHGDCNLSLDEYWK 479

RESULT 14
Q8NG65
ID Q8NG65 PRELIMINARY; PRT; 470 AA.
AC Q8NG65;
DT 01-OCT-2002 (TRENBLrel. 22, Created)
DT 01-OCT-2002 (TRENBLrel. 22, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE NIMA-related kinase 11S.
GN NEK11S.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Noguichi K., Fukazawa H., Murakami Y., Uehara Y.;
RT "Nek11, a new member of the NIMA family of kinases, involved in DNA
RT replication and genotoxic stress responses.";
RT Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AB071997; BAC06351.1; -.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0016740; F:transferase activity; IEA.
DR GO; GO:0008468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
KW ATP-binding; Kinase; Transferase.
SQ SEQUENCE 470 AA; 54006 MW; F52B82ED2B096FFB CRC64;

Query Match 16.1%; Score 583; DB 4; Length 470;
Best Local Similarity 27.9%; Pred. No. 1e-29;
Matches 155; Conservative 93; Mismatches 161; Indels 146; Gaps 15;

QY 3 KYDVLKAIQGAAGKAYLA---KGKSDSKHCVIKINFEKMPIQEKEASKKEVILLERKK 59
Db 28 RYVLQKLGSGFTVYLVSDKKARGBELKVKISVGLNPNETVQANLEAQLLSKLD 87
QY 60 HPNIVAFNFSQENGRLFIVMEYCDGGLMKRIN--RQGVLFSEDDQLGFWFQVLSGLK 117
Db 88 HPAIVKFAHFVEQDNFCIITEYCEGRDLDDKIQYKQAGKIFPENQIIEFWIQLLQVD 147
QY 118 HHDRKILHRDIKAQNIPLSKNGMVKAGLDGFIARVANNMELARTCTGTYPLSPICQ 177
Db 148 YHMERILHRDLKSNVFNKNN--LLKIGDFGVSELMLGSDCLATLTGTTHYMSPEALK 205
QY 178 NKPYNKNTDIWSLGVLYELCTLKHPPFEGNNIQQVLVKICQAHFAPISGFSGRELHSLIS 237
Db 206 HOGYDTKSDIWSLACILYEMCCWNAFAGSNFLSVLKIIVEGDTSLPERYFKELNAIME 265
QY 238 QLFQVSPDRPRDINSILKRPF----LENLIPKYLTFEVIQEEFSSMLICRAGAPASRHAG 293

```

```

Db 266 SMLNKNFSLRPSALEILKIPYLBQLQNLMCRY-----SEMTEKNDLCQKEAA 315
QY 294 KWV-----QKCKIQKVRFG---KPPRSRISVPIKRNAILFRNWRPPAGAKARSIKVI 346
Db 316 HIINAMKRIHLQTLRALSEVQKMTPRMLR-----KLQAADEKARKLKKI 363
QY 347 ERPIAAVCGHYDYVYLAQDLWLRRAHKPSVHPQPQENTGVEDYQGETRIGSPSQWPAE 406
Db 364 ----- 363
QY 407 YLQKFEAQYQYKLVKVEKQLGRFPSSAEPNYNQQLASNGEPRFQELPFRKNEKQEY 466
Db 364 -VEEKY-----EENSKRQQLASR--NFQQLSV--DVLHEKTH 396
QY 467 WKQLEETIRQQYHNDM-----KBIRKMGREPPENSKISHKTYLVKKSNIPLVHQ-----D 515
Db 397 LKGMEEKEQPEGRILSCSPQDEDERWQGREESDEP-----TLENIPESQIPSM 448
QY 516 ASEGEAPVQDIKDL 530
Db 449 LHELESIVEDATSDL 463

RESULT 15
Q8CQ04
ID Q8CQ04 PRELIMINARY; PRT; 628 AA.
AC Q8CQ04;
DT 01-MAR-2003 (TRENBLrel. 23, Created)
DT 01-MAR-2003 (TRENBLrel. 23, Last sequence update)
DT 01-OCT-2003 (TRENBLrel. 25, Last annotation update)
DE Similar to mRNA.
GN 4932416N14RIK.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RA The RIKEN Genome Exploration Research Group Phase I & II Team;
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.";
RL Nature 420:563-573(2002).
DR EMBL; AK030042; BAC26756.1; -.
DR MGD; MGI:2442276; 4932416N14RIK.
DR GO; GO:0005524; F:ATP binding; IEA.
DR GO; GO:0004674; F:protein serine/threonine kinase activity; IEA.
DR GO; GO:0004713; F:protein-tyrosine kinase activity; IEA.
DR GO; GO:0008468; P:protein amino acid phosphorylation; IEA.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR008271; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TK; 1.
DR SMART; SM00219; TYRK; 1.
DR PROSITE; PS00111; PROTEIN KINASE DOM; 1.
DR PROSITE; PS00108; PROTEIN KINASE_ST; 1.
SQ SEQUENCE 628 AA; 71666 MW; AF8672BF80DCF71 CRC64;

Query Match 16.1%; Score 583; DB 11; Length 628;
Best Local Similarity 27.9%; Pred. No. 1.5e-29;
Matches 182; Conservative 108; Mismatches 233; Indels 130; Gaps 23;

QY 3 KYDVLKAIQGAAGKAYLA---KGKSDSKHCVIKINFEKMPIQEKEASKKEVILLERKK 59
Db 29 RYVLQKLGSGFTVYLVSDKKARGBELKVKISVGLNPNETVQANLEAQLLSRLH 89
QY 60 HPNIVAFNFSQENGRLFIVMEYCDGGLMKRIN--RQGVLFSEDDQLGFWFQVLSGLK 117

```

Db 89 HPALVRFHASFNEQTCITTEYCEGRDLDYRIQYKEAGKVFAENQIVWFQILLGVD 148
QY 118 HIHDRKILHRDIKAQNIKFNKMGVAKLDFGIARVLNNSMELARTCIGTPYILSPICQ 177
Db 149 YMHERRILHRDLKSNIFLKN--LLKIGDFGSRLLMGSCELATTLTGTPHYMSPEALK 206
QY 178 NKPYNKTDIWSLGVVLCVTLKHPFEGNLLQVLKICQAHFAPISPGFSRELHSLIS 237
Db 207 HGYDAKSDIWSLACILYEMCCLHAFAGSFLSVLNVIEGKTPSPDRYPRELTIME 266
QY 238 QLFQVSPDRPSINSILKRPFLN---LIPKYLTPVIOEEFSSHMLITCRAGAPASRHAG 293
Db 267 RMLNKSPLRPSAADILKAPYMEEOQLLMCKY--PENTLEDKNS--VCQKEA--AHTI 319
QY 294 KVVQK-----CKIQVRFQCKPPRSRISVPKRNAILHRNEW--RPPAGAQAARSIXMIE 347
Db 320 NAVQKHLQTLQALSDTQKTPRERM-----WLRKQAAADERARLKKIA 365
QY 348 RPKIAVCGHYDYVYAQDMLRR---AHKPSYHPIPOENTGVEDYQETRHGSPSQW 403
Db 366 EE-----NYKENDKRMQALRSNNGSVHVAHVHLDERT--LESL-----PEQSL 409
QY 404 PA---EYLQKPE-----AQYKLVKQGLRPSSEAFNPNQCELRNSGEEPRFOELP 455
Db 410 PCLDLLEPSLEDTIVDLGHYEIPED-----PLVAEQYYSVDFDSCSDSEERQEBEMI 463
QY 456 FRK--NEMKEQB-----YWKQ-----LEEIRQQYHNDMKEI----- 484
Db 464 FSEAGDTKKEESPVYRTNQODSDTAALVGCLEHLVGYTSLDTKITNAVTMSPGMV 523
QY 485 -----RKKMGREPEEN-SKISHKTYLVKSNLPVHQDASEGEAPVQDIEKDLQOMELQN 537
Db 524 FNSAVARTKMKEMKESAVQKLGMEVF-----BEVYDYLKRRQRN 563
QY 538 TKESKNPEQKYAKKGVKFEINLDCISDENILQEEENAMDIPNETLTFEDGMK 590
Db 564 AREAEIWEHLETVVPRASDCFEVDQLYFEELLTMEGKPSLQNLPCEAQK 616

Search completed: October 8, 2004, 17:09:07
Job time : 57 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2004, 17:06:41 ; Search time 21 Seconds
(without alignments)
1679.073 Million cell updates/sec

Title: US-09-940-921B-2
Perfect score: 3616
Sequence: 1 MDKYDVIKALGQAFKAYL.....SVSRQEGTKDPYSPVLILM 683

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/2/iaa/5A COMB.psp:*
2: /cgn2_6/ptodata/2/iaa/5B COMB.psp:*
3: /cgn2_6/ptodata/2/iaa/6A COMB.psp:*
4: /cgn2_6/ptodata/2/iaa/6B COMB.psp:*
5: /cgn2_6/ptodata/2/iaa/PTUS COMB.psp:*
6: /cgn2_6/ptodata/2/iaa/backfiles1.psp:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3081.5	85.2	640	US-09-873-404-2	Sequence 2, Appli
2	1066	29.5	616	US-09-873-404-4	Sequence 4, Appli
3	816.5	22.6	345	US-09-173-581-6	Sequence 6, Appli
4	816.5	22.6	345	US-09-420-915-6	Sequence 6, Appli
5	634	17.5	692	US-09-992-481-2	Sequence 2, Appli
6	539.5	14.9	445	US-09-167-206-2	Sequence 2, Appli
7	535.5	14.8	979	US-08-870-529-2	Sequence 2, Appli
8	535.5	14.8	979	US-09-544-794-2	Sequence 2, Appli
9	517.5	14.3	699	US-09-457-040B-18	Sequence 18, Appli
10	462.5	12.8	302	US-09-221-235-2	Sequence 2, Appli
11	462.5	12.8	302	US-09-221-928-2	Sequence 2, Appli
12	462.5	12.8	302	US-09-221-527-2	Sequence 2, Appli
13	462.5	12.8	302	US-09-221-236-2	Sequence 2, Appli
14	462.5	12.8	302	US-09-221-416-2	Sequence 2, Appli
15	462.5	12.8	302	US-09-221-245-2	Sequence 2, Appli
16	462.5	12.8	302	US-09-163-115-2	Sequence 2, Appli
17	462.5	12.8	302	US-09-221-528-2	Sequence 2, Appli
18	462.5	12.8	302	US-09-593-553-2	Sequence 2, Appli
19	462.5	12.8	302	US-09-221-237-2	Sequence 2, Appli
20	458.5	12.7	302	US-09-579-664B-9	Sequence 9, Appli
21	427	11.8	233	US-08-712-709-1	Sequence 1, Appli
22	427	11.8	233	US-09-111-444-1	Sequence 1, Appli
23	427	11.8	233	US-09-541-228-1	Sequence 1, Appli
24	417.5	11.5	647	US-09-031-563-7	Sequence 7, Appli
25	417.5	11.5	647	US-09-392-277-7	Sequence 7, Appli
26	417.5	11.5	647	US-09-258-000-7	Sequence 7, Appli
27	417.5	11.5	648	US-09-031-563-5	Sequence 5, Appli

28	417.5	11.5	648	4	US-09-392-277-5	Sequence 5, Appli
29	417.5	11.5	648	4	US-09-258-000-5	Sequence 5, Appli
30	417.5	11.5	1315	3	US-09-031-563-2	Sequence 10, Appli
31	417.5	11.5	1315	4	US-09-293-505-10	Sequence 2, Appli
32	417.5	11.5	1315	4	US-09-392-277-2	Sequence 2, Appli
33	417.5	11.5	1315	4	US-09-258-000-2	Sequence 25, Appli
34	414.5	11.5	1315	3	US-09-031-563-25	Sequence 25, Appli
35	414.5	11.5	1315	4	US-09-392-277-25	Sequence 25, Appli
36	414.5	11.5	1315	4	US-09-258-000-25	Sequence 1, Appli
37	413	11.4	293	2	US-09-049-871-1	Sequence 1, Appli
38	413	11.4	293	3	US-09-295-068-1	Sequence 1, Appli
39	405.5	11.2	260	3	US-09-031-563-24	Sequence 24, Appli
40	405.5	11.2	260	4	US-09-392-277-24	Sequence 24, Appli
41	405.5	11.2	260	4	US-09-258-000-24	Sequence 24, Appli
42	403.5	11.2	925	1	US-08-252-995D-4	Sequence 4, Appli
43	403.5	11.2	925	2	US-08-834-108-4	Sequence 4, Appli
44	395.5	10.9	1269	4	US-09-645-456A-15	Sequence 15, Appli
45	395.5	10.9	1269	4	US-09-425-324A-15	Sequence 15, Appli

ALIGNMENTS

RESULT 1
US-09-873-404-2
; Sequence 2, Application US/09873404
; Patent No. 6500656
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001212-CIP
; CURRENT APPLICATION NUMBER: US/09/873,404
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2
; LENGTH: 640
; TYPE: PRT
; ORGANISM: Human
US-09-873-404-2

Query Match	85.2%	Score	3081.5	DB 4	Length	640			
Best Local Similarity	94.1%	Pred. No.	9.6e-236						
Matches	585	Conservative	2	Mismatches	4	Indels	31	Gaps	1
Qy	1	MDKYDVIKAGQAFKAYLAKGSDSKHCVIKEINFXMPIQEKASKEVILLKMKH	60						
Db	1	MDKYDVIKAGQAFKAYLAKGSDSKHCVIKEINFXMPIQEKASKEVILLKMKH	60						
Qy	61	PNIVAFNFSQENGRLFIWVEYCDGGDLMKRINRQGVLFSDQILLGWFVQISLGLKHH	120						
Db	61	PNIVAFNFSQENGRLFIWVEYCDGGDLMKRINRQGVLFSDQILLGWFVQISLGLKHH	120						
Qy	121	DRKILHRDIAQNIPLSKNGWAKLDFGIARVLNNSMELATCTGTPVYLSPEICQKP	180						
Db	121	DRKILHRDIAQNIPLSKNGWAKLDFGIARVLNNSMELATCTGTPVYLSPEICQKP	180						
Qy	181	YNNKTDIWSLGCYLYELCTLKHPFEGNNLQQLVLTICQAHFAPISPGFSRELHLSLSQLF	240						
Db	181	YNNKTDIWSLGCYLYELCTLKHPFEGNNLQQLVLTICQAHFAPISPGFSRELHLSLSQLF	240						
Qy	241	QVSPDRPDSINILKRPFFLENIIPKYLTPVEIQEEFSHMLICRAGAPASRHAGKVQCK	300						
Db	241	QVSPDRPDSINILKRPFFLENIIPKYLTPVEIQEEFSHMLICRAGAPASRHAGKVQCK	300						
Qy	301	IQKVRFGCKPPRSRISVPIKRNAILHRNEMPPAGAKARSIKMIERPKIAAVCGHYD	360						
Db	301	IQKVRFGCKPPRSRISVPIKRNAILHRNEMPPAGAKARSIKMIERPKIAAVCGHYD	360						
Qy	361	YYAQLDMLRRRAKPSYHDIPOENTGVEDYQGETRHGSPSQWPAEYLOKFEAQYK	420						

Db 361 YYAQLDMLRRRAHKPSYHPIQENTGVEDYQGETRHGSPSQWPAEYLQRFKFAQQYK 420
Qy 421 VEKQGLRPSSAEPNPNORQELRNGEPRFQELPFRKNEMKEQYKQLEELRQQYHND 480
Db 421 VEKQGLRPSSAEPNPNORQELRNGEPRFQELPFRKNEMKEQYKQLEELRQQYHND 480
Qy 481 MKEIRKXGGRPEENS KISHKTYLVKKS NLPVHQDASEGEAPVODIEKDLQKMLQNTKE 540
Db 481 MKEIRKXGGRPEENS KISHKTYLVKKS NLPVHQDASEGEAPVODIEKDLQKMLQNTKE 540
Qy 541 SKNPQKYKAKGVKFEINLDCISDENILOEEBAMDPNETLAFEGMKFEYECVKEH 600
Db 510 SKNPQKYKAKGVKFEINLDCISDENILOEEBAMDPNETLAFEGMKFEYECVKEH 569
Qy 601 GDYTDKAFKHLCPAGFSTOT 622
Db 570 GDYTDKAFKHLCPAATLT 591
RESULT 2
US-09-873-404-4
; Sequence 4, Application US/09873404
; Patent No. 6500656
; GENERAL INFORMATION:
; APPLICANT: WEBSTER, Marion et al
; TITLE OF INVENTION: ISOLATED HUMAN KINASE PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN KINASE PROTEINS, AND USES
; FILE REFERENCE: CL001212-CIP
; CURRENT APPLICATION NUMBER: US/09/873,404
; CURRENT FILING DATE: 2001-06-05
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 4
; LENGTH: 616
; TYPE: PRT
; ORGANISM: Mouse
US-09-873-404-4
Query Match 29.5%; Score 1066; DB 4; Length 616;
Best Local Similarity 38.7%; Pred. No. 4e-76;
Matches 245; Conservative 95; Mismatches 147; Indels 146; Gaps 16;
Qy 1 MDKYDVKAIGGAGFKAYLAKGSDSKHCYVKEINFEKMPIOEKEASKEVILLKMKH 60
Db 1 MEKYVRLQKIGSGFKAVLVKSTEDGRHYVKEINISMSKDEESRREAVLANMKH 60
Qy 61 PNIVAFNSFOENGRLFIWMEYCDGDLMKRINRQGVLFSEDLQILGWVQISLGLKH 120
Db 61 PNIVQKESFENGSLYIWMYCEGDLFKRINAQKGFALFOEDQILDWVQICLALKH 120
Qy 121 DRKILHRDIKAQNIPLSKNGVMVAKLGDGFIARVLNNSMELARTCTGTPYLSPEICONKP 180
Db 121 DRKILHRDIKSNIFLTKDGTGTV-QLGDFGIARVLNSTVELARTCTGTPYLSPEICENKP 179
Qy 181 YNNKTDIWSLGCVLVELCTLKHFFEGNNLQQLVLKICQAHFAPISPGFSRELHSLISOLF 240
Db 180 YNNKSDIWAIGCVLYELCTLKHAFEGAGNMKNLVKIIISGSPPPVSHYSYDLRSLLSOLF 239
Qy 241 QVSPDRPSINSILKRPFLNIPKVLTPVTEQEFESHMLICRAG----- 285
Db 240 KRNPRDRPSVNSILEKGFIAKIEFLSPQLIAEEFCLTKTSKFGPQLPKRPSAGGV 299
Qy 286 -----APASRHAGVQKCIQKVRFGCKPPRSRIS--VPKRNAILHRNWRP 333
Db 300 SSFVPAQKITKPAKYGVPLTKYKGDKKLEKPKPKHQAQHPVKK--MNSGERK 356
Qy 334 PAG--AQKARSIMTER----- 348
Db 357 KMSSEAAKRRLEFIEKEKKQKDQIRFLKABQMKFEQKQRLERINRAREQQWRNVLRAG 416
Qy 349 -----PKIAAVCGHYDYVYAAQLDMLRR-RAH-----KPSYHF--IPQE 383

Db 417 SEVKASFFGIGAVSPSPGSPRGQYEHYHAIFDQMLRAEDNEARWKGLIYGNLPER 476
Qy 384 NTG---VEDYQGETRHGSPSQWPAEYLQRFKFAQQYKLVKVE-----KQLGLRPSS 431
Db 477 QKGLHAEVRAHQ-----VEEFLQRKEAMQKARAEHGVVYLARLQRILQ--- 522
Qy 432 APNPNORQELR-----SNGEPRFQELPFRKNEMKEQYKQLEELRQQYH 479
Db 523 ---NFNERQKIKAKLRGENKEADGKGQFAT--BETDMR---LKKWESLKAQTNARAVALK 575
Qy 480 DMKEIRKXGGRPEENS KISHKTYLVKKS NLPV 512
Db 576 EQLEKRRKAEYEREKKVWEHLVARVKSDVPL 608
RESULT 3
US-09-173-581-6
; Sequence 6, Application US/09173581A
; Patent No. 6013455
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Tang, Y. Tom
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Yue, Henry
; APPLICANT: Guegler, Karl J.
; APPLICANT: Corley, Neil C.
; APPLICANT: Gorgone, Gina
; APPLICANT: Azimzai, Yaida
; APPLICANT: Lu, Aina
; TITLE OF INVENTION: Protein Kinase Homologs
; FILE REFERENCE: PF-0614 US
; CURRENT APPLICATION NUMBER: US/09/173,581A
; CURRENT FILING DATE: 1998-10-15
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PERL Program
; SEQ ID NO 6
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: 1567782
US-09-173-581-6
Query Match 22.6%; Score 816.5; DB 3; Length 345;
Best Local Similarity 44.9%; Pred. No. 1e-56;
Matches 173; Conservative 55; Mismatches 80; Indels 77; Gaps 6;
Qy 1 MDKYDVKAIGGAGFKAYLAKGSDSKHCYVKEINFEKMPIOEKEASKEVILLKMKH 60
Db 1 MEKYVRLQKIGSGFKAVLVKSTEDGRHYVKEINISMSKDEESRREAVLANMKH 60
Qy 61 PNIVAFNSFOENGRLFIWMEYCDGDLMKRINRQGVLFSEDLQILGWVQISLGLKH 120
Db 61 PNIVQYRESF-----EGILDWVQICLALKH 88
Qy 121 DRKILHRDIKAQNIPLSKNGVMVAKLGDGFIARVLNNSMELARTCTGTPYLSPEICONKP 180
Db 89 DRKILHRDIKSNIFLTKDGTGTV-QLGDFGIARVLNSTVELARTCTGTPYLSPEICENKP 147
Qy 181 YNNKTDIWSLGCVLVELCTLKHFFEGNNLQQLVLKICQAHFAPISPGFSRELHSLISOLF 240
Db 148 YNNKSDIWAIGCVLYELCTLKHAFEGAGNMKNLVKIIISGSPPPVSHYSYDLRSLLSOLF 207
Qy 241 QVSPDRPSINSILKRPFLNIPKVLTPVTEQEFESHMLICRAG---PASSHAG--- 293
Db 208 KRNPRDRPSVNSILEKGFIAKIEFLSPQLIAEEFCLTKTSKFGPQLPKRPSAGGV 267
Qy 294 -----KVQCKCIQKVRFGCKPPRSRISVPKRNAIL 325
Db 268 ISVMPAQKITKPAKYGVPLTKYKGDKKLEKPKPKHQAQHPVKK--MNSGERK 326
Qy 326 LHRNWRPAPAGAKARSIMTERPK 350

Query Match 14.9%; Score 539.5; DB 4; Length 445;
Best Local Similarity 29.9%; Pred. No. 1.3e-34;
Matches 141: Conservative 88; Mismatches 167; Indels 75; Gaps 13;

```

QY      2  DKYDVWIKAGOGAFGKAYLAKGKSDSKHCWIKIEINFKMPIQBEAKSKVEVILLKMKHP 61
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      6  EDYEVLYTIGTSGYRCOKIRKSDGKILVWKELDYGSMTEABQKMLVSEVNLRELKHP 65
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY     62  NTVAFNFSF--QENGRLPIVMEYCGDGLMKRI---NRQGVLFVSEDOILGMFWOISLGL 116
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db     66  NIVRYDYRIIDRTNTLLIVMEYCEGDDLAVITKTKERQYL--DEEFVLVWMTQLTAL 124
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    117  KXIHPRK-----ILHRDIKAONIFLSKXNGWAKLGDPGFIARVUNNMEIARTCIGTPYVL 171
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    125  KECHRRSDGGHTVLHRDLKPAVFLDQKNV-KLGDPLGRLAIINHDTSPAKTFVGTPTYM 183
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    172  SPEI CONKPYNNKTDIWSLGVLYELCTLKHPPPEGNNLQOLVKI COAHFAPISPGFSRE 231
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    184  SPEQNNRMSYNEKSDIWSLGLLYELCALMPPTAFSQKELAGIKREGKFRPIRYRSD 243
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    232  LHSLSIQLFQVSPDRPDSINSILKRPFLNLI PKYLTPEVIQEFSHMLICRA---GAP- 287
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    244  LNEIITRMLNLKHVRPSVESILENPJADL-----VADEQRNLERRGOLGSP 294
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    288  ASRHAGKVQKQIKVRFQCKPGRSRI SVPKRNAILHRNWRPAPGAQK----- 339
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    295  KSQDSSPVLSELKKEIQLO-----ERERALKARERLEQKEQLCVRERLA 341
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    340  -----ARSIKWIERPIAAVCGHYDYVAQLDMLPRRAHKSHYHPIQENTGVED 389
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    342  EDKLARENLKNVSLKERFLSLANPELLNPSSVIKKVH-----FSGESKENI---- 394
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY    390  YQGETRHGFPSPQWPAEYLQKFBAAQ-----YKLVKEQKOLGR 428
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db    395  MRSENSEBQLTSKSKCOLKRLHAAOLRAQALSDIEKNYQKSLQILGWR 445
      : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 7

US-08-870-529-2

```

Sequence 2, Application US/08870529
Patent No. 6080557
GENERAL INFORMATION:
APPLICANT: Sims, John E.
APPLICANT: Virca, G. Duke
APPLICANT: Bird, Timothy A.
APPLICANT: Anderson, Dirk M.
TITLE OF INVENTION: IL-1/INF-(-ACTIVATED KINASE (ITAK),
NUMBER OF SEQUENCES: 18
CORRESPONDENCE ADDRESS:
ADDRESSEE: SEED AND BERRY LLP
STREET: 6300 Columbia Center, 701 Fifth Avenue
CITY: Seattle
STATE: Washington
COUNTRY: USA
ZIP: 98104
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/870,529
FILING DATE: 06-JUN-1997
CLASSIFICATION: 435
ATTORNEY/AGENT INFORMATION:
NAME: Mcmasters, David D.
REGISTRATION NUMBER: 33,963
REFERENCE/DOCKET NUMBER: 480052.418
TELECOMMUNICATION INFORMATION:
TELEPHONE: (206) 622-4900
TELEFAX: (206) 682-6031
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 979 amino acids
TYPE: amino acid

```

```

;
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-870-529-2

Query Match      14.8%; Score 535.5; DB 3; Length 979;
Best Local Similarity 40.7%; Pred. NO. 7.8e-34;
Matches 105; Conservative 55; Mismatches 95; Indels 3; Gaps 2;

QY      4 YDVIRAIQAGAFGKAYLAGKSGDSKHCVTKEINFEKMPIQEAKSKEVILLKMKHPNI 63
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      52 YPIRVLGAGAGATLTARTEDDSILVWKEVDLTLFSEKRRDALNEIVILALICHONI 111

QY      64 VAFNFSFOENGLFTVMEYCDGDLMKRINRQGVLFSEDIQLHFWQVSLGKIHDRK 123
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      112 IAYVNHFMNTTLLILEYCNCGNLYDKILRKQDKLFEEMVWVLFQVSAVSTIHKAG 171

QY      124 ILHRDIKQANI ELSKNGWAKLGDGFIARVLANSMELATCTGTPYLSPEICQKNPYN 183
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      172 ILHRDIKTNI EFTKANLI-KUGDYGKLAKLNSYMSMAETLVGTPYMSPELCQGVKNF 230

QY      184 KTDIVSLGCVLYELCTLKHPPFEGNNLQQIVLXICQAHFA--DISPGFSRSLHSLSQLFQ 241
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      231 KSDIVWAGCVIFELLTKRTFDATPNLNCVKIVQIRAMEVDSQYSLELQWVHSCLD 290

QY      242 VSPDRRPSINSILKRPF 259
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db      291 QDPEQRPTADELLDRPLL 308
      : : : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 8

US-09-544-794-2

```

/ Sequence 2, Application US/09S44794
/
/ GENERAL INFORMATION:
/
/ APPLICANT: Sims, John E.
/ APPLICANT: Virca, G. Duke
/ APPLICANT: Bird, Timothy A.
/ APPLICANT: Anderson, Dirk M.
/
/ TITLE OF INVENTION: Polypeptides Having Kinase Activity
/
/ FILE REFERENCE: 2005-B
/
/ CURRENT APPLICATION NUMBER: US/09/544,794
/
/ CURRENT FILING DATE: 2000-04-07
/
/ PRIOR APPLICATION NUMBER: 08/870,529
/
/ PRIOR FILING DATE: 1997-06-06
/
/ PRIOR APPLICATION NUMBER: 60/059,979
/
/ PRIOR FILING DATE: 1996-06-10
/
/ NUMBER OF SEQ ID NOS: 18
/
/ SOFTWARE: PatentIn ver. 2.0
/
/ SEQ ID NO 2
/
/ LENGTH: 979
/
/ TYPE: PRT
/
/ ORGANISM: Homo sapiens
/
/ US-09-544-794-2

```

```

QY 621 QTVAAVGNRQWDGGAPQILLQMAVADITSTCTGPDDESVLVSROEKTQDPY 676
DB 554 QPPIFFAN-----PSRLPQMSA-----PDV-----RESKRSRPH 582

RESULT 10
US-09-221-235-2
; Sequence 2, Application US/09221235
; Patent No. 6043040
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,235
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-235-2

Query Match 12.8%; Score 462.5; DB 3; Length 302;
Best Local Similarity 37.1%; Pred. No. 9.3e-29;
Matches 95; Conservative 54; Mismatches 100; Indels 7; Gaps 5;

QY 4 YDVIKAIGOGAFGKAYLAKGDSKHCVIKBIN--FEKMPIQEKBAKKEVILLEKMKHPN 62
DB 34 FRIEKKIGRGQFSEVYRAACLLDGVVPALKKVGIFDLMDAKARADCICEIDLLKQLNHPN 93
QY 63 IVAFNFSFOENGRLFIVMEYCDGDLMKRIN--RQRGVLFSEDOILLGWFYQISLGLKXIH 120
DB 94 VKIYASFIEDNELNIVLEADAGDLSRMKKPKOKRLLIPERTVWKYFVOLCSALEHHM 153
QY 121 DRKILHRDIAQNIIFLSKNGMVAKLGDGFIARVLNNSMELARTCIGTPPYLSPEICQNKP 180
DB 154 SREVMHRDIKANVFITATGVV--KLGDGLGRFFSSKTTAAHSLVGTTPPYMSPERIENG 212
QY 181 YNKTIDWSLGCVLVYELCTIKHPFEGN--NLQOLVTKIQAHFAPI--SPGFSRELHSLIS 237
DB 213 YNFKSDIWSLGCULLYEMAAQSPFYGDKNLYSLCKKIEQCDDYPPLPDSHYSELRLVN 272
QY 238 QLPQVSPDRDRPSINSI 253
DB 273 MCINPDPKRPDVTYV 288

RESULT 11
US-09-221-928-2
; Sequence 2, Application US/09221928
; Patent No. 6131030
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,928
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE:
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-221-928-2

Query Match 12.8%; Score 462.5; DB 3; Length 302;
Best Local Similarity 37.1%; Pred. No. 9.3e-29;
Matches 95; Conservative 54; Mismatches 100; Indels 7; Gaps 5;

```

```

; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,236
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-221-236-2

Query Match      12.8%; Score 462.5; DB 3; Length 302;
Best Local Similarity 37.1%; Pred. No. 9.3e-29;
Matches 95; Conservative 54; Mismatches 100; Indels 7; Gaps 5;

QY 4 YDVIKAIQGGAFGKAYLAKGSDSKHCIVKEIN--FEKMPIQEKASKEVILLKMKHPN 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 FRIEKKIGRGQFSEVYRAACLLDGVPAVKVQVIFDLMDAKARADCICEIDLLKQLNHPN 93

QY 63 IVAFNFSQENGRLFIYMEYCDGGDLMKRIN--RQGVLFSEDOILGWFWQISLGLKHIIH 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 94 VIKYASFTIEDNELNIVLELADAGDLSRMKHFKKQKRLIPERTVWKYFVQLCSALEHMH 153

QY 121 DRKILHRDIKAQNIPLSKNGVMVAKLGDGFIARVLNNSMELARTCIGTPYVLSPEICQNK 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 154 SRVWHRDIKANVFITATGVV--KLGDGLGRFFSSKTTAAHSLVGTPTYWSPERIENG 212

QY 181 YNNKTDIWSLGCVLVELCTLKHPPEGN--NLQQLVLKICQAHFAP1-SPGFSRELHSLIS 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 YNFKSDIWSLGCLLYEMAALQSPFYGDQWNLVSLCKIEQCDYPLPSDHYSEELRLQVN 272

QY 238 QLFQVSPDRPSINSI 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 MCINPDPEKRPDVTYV 288

RESULT 14
US-09-221-416-2
; Sequence 2, Application US/09221416
; Patent No. 6153417
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
; TITLE OF INVENTION: NOVEL CSAPK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR
; FILE REFERENCE: MNI-050
; CURRENT APPLICATION NUMBER: US/09/221,416
; CURRENT FILING DATE: 1998-12-28
; EARLIER APPLICATION NUMBER: 09/163,115
; EARLIER FILING DATE: 1998-09-29
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 2
; LENGTH: 302
; TYPE: PRT
; ORGANISM: Homo sapiens
; ORGANISM: Homo sapiens
US-09-221-416-2

Query Match      12.8%; Score 462.5; DB 3; Length 302;
Best Local Similarity 37.1%; Pred. No. 9.3e-29;
Matches 95; Conservative 54; Mismatches 100; Indels 7; Gaps 5;

QY 4 YDVIKAIQGGAFGKAYLAKGSDSKHCIVKEIN--FEKMPIQEKASKEVILLKMKHPN 62
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 34 FRIEKKIGRGQFSEVYRAACLLDGVPAVKVQVIFDLMDAKARADCICEIDLLKQLNHPN 93

QY 63 IVAFNFSQENGRLFIYMEYCDGGDLMKRIN--RQGVLFSEDOILGWFWQISLGLKHIIH 120
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 94 VIKYASFTIEDNELNIVLELADAGDLSRMKHFKKQKRLIPERTVWKYFVQLCSALEHMH 153

QY 121 DRKILHRDIKAQNIPLSKNGVMVAKLGDGFIARVLNNSMELARTCIGTPYVLSPEICQNK 180
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 154 SRVWHRDIKANVFITATGVV--KLGDGLGRFFSSKTTAAHSLVGTPTYWSPERIENG 212

QY 181 YNNKTDIWSLGCVLVELCTLKHPPEGN--NLQQLVLKICQAHFAP1-SPGFSRELHSLIS 237
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 213 YNFKSDIWSLGCLLYEMAALQSPFYGDQWNLVSLCKIEQCDYPLPSDHYSEELRLQVN 272

QY 238 QLFQVSPDRPSINSI 253
   : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 273 MCINPDPEKRPDVTYV 288

RESULT 13
US-09-221-236-2
; Sequence 2, Application US/09221236
; Patent No. 6146841
; GENERAL INFORMATION:
; APPLICANT: Acton, Susan
```


Db 154 SRVVRDIPKANVFITATGVV-KLGDGLGRFFSSKTTAAHSLVGTPTYMSPERIHENG 212

QY 181 YNKTIDINSLGCVLVELCTLKHFFEGN--NLQOLVKICQAHFAP1-SPGFSRELHSLIS 237

Db 213 YNFKSDINSLGCVLVELCTLKHFFEGN--NLQOLVKICQAHFAP1-SPGFSRELHSLIS 237

QY 238 QLFQVSPDRDRPSINSI 253

Db 273 MCINPDPEKRPDVTYV 288

RESULT 15

US-09-221-245-2

Sequence 2, Application US/09221245

Patent No. 6180358

GENERAL INFORMATION:

APPLICANT: Acton, Susan

TITLE OF INVENTION: NOVEL CSAFK-1 NUCLEIC ACID MOLECULES AND USES THEREFOR

FILE REFERENCE: MNI-050

CURRENT APPLICATION NUMBER: US/09/221,245

CURRENT FILING DATE: 1998-12-28

EARLIER APPLICATION NUMBER: US 09/163,115

EARLIER FILING DATE: 1998-09-29

NUMBER OF SEQ ID NOS: 15

SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 2

LENGTH: 302

TYPE: PRT

ORGANISM: Homo sapiens

US-09-221-245-2

Query Match 12.8%; Score 462.5; DB 3; Length 302;

Best Local Similarity 37.1%; Pred. No. 9.3e-29;

Matches 95; Conservative 54; Mismatches 100; Indels 7; Gaps 5;

QY 4 YDVKAIGQAGFKAYLAKGSDSKHCVIKEIN-FERMPIOKEASKKEVILLEKMKHPN 62

Db 34 FRIEKKIGRGQSEVYRAACLLDGVFVALKQVIFDLMDAKARADCICEIDLLKOLNHPN 93

QY 63 IVAFNSFOENGRLEFTVMEYCDGGLMKRIN--RQGVLFSEDOILGNFVQISLGLKHIH 120

Db 94 VIKYVASFLEDNELNIVLELADAGDLSRWIKHFKKQKLIPIERTVWKYFVQLCSALEHMH 153

QY 121 DRKILHRDIKAQNIPLSKNGMVAKLGDGFIARVLNNSMELARTCIGTPPYLSPEICQNKP 180

Db 154 SRVVRDIPKANVFITATGVV-KLGDGLGRFFSSKTTAAHSLVGTPTYMSPERIHENG 212

QY 181 YNKTIDINSLGCVLVELCTLKHFFEGN--NLQOLVKICQAHFAP1-SPGFSRELHSLIS 237

Db 213 YNFKSDINSLGCVLVELCTLKHFFEGN--NLQOLVKICQAHFAP1-SPGFSRELHSLIS 237

QY 238 QLFQVSPDRDRPSINSI 253

Db 273 MCINPDPEKRPDVTYV 288

Search completed: October 8, 2004, 17:09:58

Job time : 24 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2004, 17:01:35 ; Search time 15 Seconds
(without alignments)
2370.928 Million cell updates/sec

Title: US-09-940-921B-2
Perfect score: 3616
Sequence: 1 MDKVDVKAIGCGAGKAYL.....SVSRQEKTKDPYSPVLILM 683

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwisProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1106.5	30.6	774	1 NEK1 MOUSE	P51954 mus musculus
2	1063.5	29.4	1258	1 NEK1 HUMAN	Q96py6 homo sapien
3	808.5	22.4	511	1 NEK3 MOUSE	Q9r0a5 mus musculus
4	808	22.3	506	1 NEK3 HUMAN	P51956 homo sapien
5	665.5	18.4	792	1 NEK4 MOUSE	Q921j2 mus musculus
6	660	18.3	841	1 NEK4 HUMAN	P51957 homo sapien
7	634	17.5	692	1 NEK8 HUMAN	Q858g6 homo sapien
8	633	17.5	698	1 NEK8 MOUSE	Q921r4 mus musculus
9	623	17.2	697	1 NEK8 BRARE	Q90xc2 brachydanio
10	539.5	14.9	445	1 NEK2 HUMAN	P51955 homo sapien
11	535.5	14.8	979	1 NEK2 MOUSE	Q8td19 homo sapien
12	532.5	14.7	443	1 NEK2 HUMAN	O35942 mus musculus
13	532.5	14.7	984	1 NEK9 MOUSE	Q8k1r7 mus musculus
14	524.5	14.5	944	1 NEK9 XENLA	Q72zc8 xenopus lae
15	521.5	14.4	431	1 NRKA TRYBB	Q08942 trypanosoma
16	517.5	14.3	699	1 NIMA EMENI	P11837 emericella
17	511	14.1	722	1 FIN1 SCHPO	O13839 schizosacch
18	490	13.6	431	1 NRK8 TRYBB	Q03428 trypanosoma
19	472.5	13.1	313	1 NEK6 RAT	P59895 rattus norv
20	471.5	13.0	313	1 NEK6 HUMAN	Q9hc98 homo sapien
21	469.5	13.0	213	1 NEK6 MOUSE	Q9s870 mus musculus
22	464.5	12.8	779	1 NIMI NEUCR	P48479 neurospora
23	462.5	12.8	302	1 NEK7 HUMAN	Q8tdx7 homo sapien
24	458.5	12.7	302	1 NEK7 MOUSE	Q9es74 mus musculus
25	423	11.7	435	1 KIN3 YEAST	P22209 saccharomyc
26	405	11.2	576	1 POLO DROME	P52304 drosophila
27	402.5	11.1	1576	1 YLK3 CAEEL	P41951 caenorhabdi
28	395	10.9	1360	1 TNK1 HUMAN	Q9uke5 homo sapien
29	392.5	10.9	966	1 STKA MOUSE	O55098 mus musculus
30	392	10.8	1142	1 GIN4 YEAST	Q12263 saccharomyc
31	391	10.8	658	1 PAK1 SCHPO	P50527 schizosacch
32	389	10.8	968	1 STKA HUMAN	O94804 homo sapien
33	388	10.7	705	1 CDC5 YEAST	P32562 saccharomyc

us-09-940-921b-2.rsp

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2004, 17:01:35 ; Search time 15 Seconds
(without alignments)
2370.928 Million cell updates/sec

Title: US-09-940-921B-2
Perfect score: 3616
Sequence: 1 MDKVDVKAIGCGAGKAYL.....SVSRQEKTKDPYSPVLILM 683

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwisProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1106.5	30.6	774	1 NEK1 MOUSE	P51954 mus musculus
2	1063.5	29.4	1258	1 NEK1 HUMAN	Q96py6 homo sapien
3	808.5	22.4	511	1 NEK3 MOUSE	Q9r0a5 mus musculus
4	808	22.3	506	1 NEK3 HUMAN	P51956 homo sapien
5	665.5	18.4	792	1 NEK4 MOUSE	Q921j2 mus musculus
6	660	18.3	841	1 NEK4 HUMAN	P51957 homo sapien
7	634	17.5	692	1 NEK8 HUMAN	Q858g6 homo sapien
8	633	17.5	698	1 NEK8 MOUSE	Q921r4 mus musculus
9	623	17.2	697	1 NEK8 BRARE	Q90xc2 brachydanio
10	539.5	14.9	445	1 NEK2 HUMAN	P51955 homo sapien
11	535.5	14.8	979	1 NEK2 MOUSE	Q8td19 homo sapien
12	532.5	14.7	443	1 NEK2 HUMAN	O35942 mus musculus
13	532.5	14.7	984	1 NEK9 MOUSE	Q8k1r7 mus musculus
14	524.5	14.5	944	1 NEK9 XENLA	Q72zc8 xenopus lae
15	521.5	14.4	431	1 NRKA TRYBB	Q08942 trypanosoma
16	517.5	14.3	699	1 NIMA EMENI	P11837 emericella
17	511	14.1	722	1 FIN1 SCHPO	O13839 schizosacch
18	490	13.6	431	1 NRK8 TRYBB	Q03428 trypanosoma
19	472.5	13.1	313	1 NEK6 RAT	P59895 rattus norv
20	471.5	13.0	313	1 NEK6 HUMAN	Q9hc98 homo sapien
21	469.5	13.0	213	1 NEK6 MOUSE	Q9s870 mus musculus
22	464.5	12.8	779	1 NIMI NEUCR	P48479 neurospora
23	462.5	12.8	302	1 NEK7 HUMAN	Q8tdx7 homo sapien
24	458.5	12.7	302	1 NEK7 MOUSE	Q9es74 mus musculus
25	423	11.7	435	1 KIN3 YEAST	P22209 saccharomyc
26	405	11.2	576	1 POLO DROME	P52304 drosophila
27	402.5	11.1	1576	1 YLK3 CAEEL	P41951 caenorhabdi
28	395	10.9	1360	1 TNK1 HUMAN	Q9uke5 homo sapien
29	392.5	10.9	966	1 STKA MOUSE	O55098 mus musculus
30	392	10.8	1142	1 GIN4 YEAST	Q12263 saccharomyc
31	391	10.8	658	1 PAK1 SCHPO	P50527 schizosacch
32	389	10.8	968	1 STKA HUMAN	O94804 homo sapien
33	388	10.7	705	1 CDC5 YEAST	P32562 saccharomyc

us-09-940-921b-2.rsp

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2004, 17:01:35 ; Search time 15 Seconds
(without alignments)
2370.928 Million cell updates/sec

Title: US-09-940-921B-2
Perfect score: 3616
Sequence: 1 MDKVDVKAIGCGAGKAYL.....SVSRQEKTKDPYSPVLILM 683

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwisProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1106.5	30.6	774	1 NEK1 MOUSE	P51954 mus musculus
2	1063.5	29.4	1258	1 NEK1 HUMAN	Q96py6 homo sapien
3	808.5	22.4	511	1 NEK3 MOUSE	Q9r0a5 mus musculus
4	808	22.3	506	1 NEK3 HUMAN	P51956 homo sapien
5	665.5	18.4	792	1 NEK4 MOUSE	Q921j2 mus musculus
6	660	18.3	841	1 NEK4 HUMAN	P51957 homo sapien
7	634	17.5	692	1 NEK8 HUMAN	Q858g6 homo sapien
8	633	17.5	698	1 NEK8 MOUSE	Q921r4 mus musculus
9	623	17.2	697	1 NEK8 BRARE	Q90xc2 brachydanio
10	539.5	14.9	445	1 NEK2 HUMAN	P51955 homo sapien
11	535.5	14.8	979	1 NEK2 MOUSE	Q8td19 homo sapien
12	532.5	14.7	443	1 NEK2 HUMAN	O35942 mus musculus
13	532.5	14.7	984	1 NEK9 MOUSE	Q8k1r7 mus musculus
14	524.5	14.5	944	1 NEK9 XENLA	Q72zc8 xenopus lae
15	521.5	14.4	431	1 NRKA TRYBB	Q08942 trypanosoma
16	517.5	14.3	699	1 NIMA EMENI	P11837 emericella
17	511	14.1	722	1 FIN1 SCHPO	O13839 schizosacch
18	490	13.6	431	1 NRK8 TRYBB	Q03428 trypanosoma
19	472.5	13.1	313	1 NEK6 RAT	P59895 rattus norv
20	471.5	13.0	313	1 NEK6 HUMAN	Q9hc98 homo sapien
21	469.5	13.0	213	1 NEK6 MOUSE	Q9s870 mus musculus
22	464.5	12.8	779	1 NIMI NEUCR	P48479 neurospora
23	462.5	12.8	302	1 NEK7 HUMAN	Q8tdx7 homo sapien
24	458.5	12.7	302	1 NEK7 MOUSE	Q9es74 mus musculus
25	423	11.7	435	1 KIN3 YEAST	P22209 saccharomyc
26	405	11.2	576	1 POLO DROME	P52304 drosophila
27	402.5	11.1	1576	1 YLK3 CAEEL	P41951 caenorhabdi
28	395	10.9	1360	1 TNK1 HUMAN	Q9uke5 homo sapien
29	392.5	10.9	966	1 STKA MOUSE	O55098 mus musculus
30	392	10.8	1142	1 GIN4 YEAST	Q12263 saccharomyc
31	391	10.8	658	1 PAK1 SCHPO	P50527 schizosacch
32	389	10.8	968	1 STKA HUMAN	O94804 homo sapien
33	388	10.7	705	1 CDC5 YEAST	P32562 saccharomyc

us-09-940-921b-2.rsp

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2004, 17:01:35 ; Search time 15 Seconds
(without alignments)
2370.928 Million cell updates/sec

Title: US-09-940-921B-2
Perfect score: 3616
Sequence: 1 MDKVDVKAIGCGAGKAYL.....SVSRQEKTKDPYSPVLILM 683

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwisProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1106.5	30.6	774	1 NEK1 MOUSE	P51954 mus musculus
2	1063.5	29.4	1258	1 NEK1 HUMAN	Q96py6 homo sapien
3	808.5	22.4	511	1 NEK3 MOUSE	Q9r0a5 mus musculus
4	808	22.3	506	1 NEK3 HUMAN	P51956 homo sapien
5	665.5	18.4	792	1 NEK4 MOUSE	Q921j2 mus musculus
6	660	18.3	841	1 NEK4 HUMAN	P51957 homo sapien
7	634	17.5	692	1 NEK8 HUMAN	Q858g6 homo sapien
8	633	17.5	698	1 NEK8 MOUSE	Q921r4 mus musculus
9	623	17.2	697	1 NEK8 BRARE	Q90xc2 brachydanio
10	539.5	14.9	445	1 NEK2 HUMAN	P51955 homo sapien
11	535.5	14.8	979	1 NEK2 MOUSE	Q8td19 homo sapien
12	532.5	14.7	443	1 NEK2 HUMAN	O35942 mus musculus
13	532.5	14.7	984	1 NEK9 MOUSE	Q8k1r7 mus musculus
14	524.5	14.5	944	1 NEK9 XENLA	Q72zc8 xenopus lae
15	521.5	14.4	431	1 NRKA TRYBB	Q08942 trypanosoma
16	517.5	14.3	699	1 NIMA EMENI	P11837 emericella
17	511	14.1	722	1 FIN1 SCHPO	O13839 schizosacch
18	490	13.6	431	1 NRK8 TRYBB	Q03428 trypanosoma
19	472.5	13.1	313	1 NEK6 RAT	P59895 rattus norv
20	471.5	13.0	313	1 NEK6 HUMAN	Q9hc98 homo sapien
21	469.5	13.0	213	1 NEK6 MOUSE	Q9s870 mus musculus
22	464.5	12.8	779	1 NIMI NEUCR	P48479 neurospora
23	462.5	12.8	302	1 NEK7 HUMAN	Q8tdx7 homo sapien
24	458.5	12.7	302	1 NEK7 MOUSE	Q9es74 mus musculus
25	423	11.7	435	1 KIN3 YEAST	P22209 saccharomyc
26	405	11.2	576	1 POLO DROME	P52304 drosophila
27	402.5	11.1	1576	1 YLK3 CAEEL	P41951 caenorhabdi
28	395	10.9	1360	1 TNK1 HUMAN	Q9uke5 homo sapien
29	392.5	10.9	966	1 STKA MOUSE	O55098 mus musculus
30	392	10.8	1142	1 GIN4 YEAST	Q12263 saccharomyc
31	391	10.8	658	1 PAK1 SCHPO	P50527 schizosacch
32	389	10.8	968	1 STKA HUMAN	O94804 homo sapien
33	388	10.7	705	1 CDC5 YEAST	P32562 saccharomyc

us-09-940-921b-2.rsp

GenCore version 5.1.6
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2004, 17:01:35 ; Search time 15 Seconds
(without alignments)
2370.928 Million cell updates/sec

Title: US-09-940-921B-2
Perfect score: 3616
Sequence: 1 MDKVDVKAIGCGAGKAYL.....SVSRQEKTKDPYSPVLILM 683

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SwisProt_42:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1106.5	30.6	774	1 NEK1 MOUSE	P51954 mus musculus
2	1063.5	29.4	1258	1 NEK1 HUMAN	Q96py6 homo sapien
3	808.5	22.4	511	1 NEK3 MOUSE	Q9r0a5 mus musculus
4	808	22.3	506	1 NEK3 HUMAN	P51956 homo sapien
5	665.5	18.4	792	1 NEK4 MOUSE	Q921j2 mus musculus
6	660	18.3	841	1 NEK4 HUMAN	P51957 homo sapien
7	634	17.5	692	1 NEK8 HUMAN	Q858g6 homo sapien
8	633	17.5	698	1 NEK8 MOUSE	Q921r4 mus musculus
9	623	17.2	697	1 NEK8 BRARE	Q90xc2 brachydanio
10	539.5	14.9	445	1 NEK2 HUMAN	P51955 homo sapien
11	535.5	14.8	979	1 NEK2 MOUSE	Q8td19 homo sapien
12	532.5	14.7	443	1 NEK2 HUMAN	O35942 mus musculus
13	532.5	14.7	984	1 NEK9 MOUSE	Q8k1r7 mus musculus
14	524.5	14.5	944	1 NEK9 XENLA	Q72zc8 xenopus lae
15	521.5	14.4	431	1 NRKA TRYBB	Q08942 trypanosoma
16	517.5	14.3	699	1 NIMA EMENI	P11837 emericella
17	511	14.1	722	1 FIN1 SCHPO	O13839 schizosacch
18	490	13.6	431	1 NRK8 TRYBB	Q03428 trypanosoma
19	472.5	13.1	313	1 NEK6 RAT	P59895 rattus norv
20	471.5	13.0	313	1 NEK6 HUMAN	Q9hc98 homo sapien
21	469.5	13.0	213	1 NEK6 MOUSE	Q9s870 mus musculus
22	464.5	12.8	779	1 NIMI NEUCR	P48479 neurospora
23	462.5	12.8	302	1 NEK7 HUMAN	Q8tdx7 homo sapien
24	458.5	12.7	302	1 NEK7 MOUSE	Q9es74 mus musculus
25	423	11.7	435	1 KIN3 YEAST	P22209 saccharomyc
26	405	11.2	576	1 POLO DROME	P52304 drosophila
27	402.5	11.1	1576	1 YLK3 CAEEL	P41951 caenorhabdi
28	395	10.9	1360	1 TNK1 HUMAN	Q9uke5 homo sapien
29	392.5	10.9	966	1 STKA MOUSE	O55098 mus musculus
30	392	10.8	1142	1 GIN4 YEAST	Q12263 saccharomyc
31	391	10.8	658	1 PAK1 SCHPO	P50527 schizosacch
32	389	10.8	968		

```

DB          730 SFSAT 734

RESULT 2
NEKI_HUMAN
ID          NEKI_HUMAN      STANDARD;      PRT; 1258 AA.
AC          Q96PY6; Q9Y594;
DT          28-FEB-2003 (Rel. 41, Created)
DT          28-FEB-2003 (Rel. 41, Last sequence update)
DT          10-OCT-2003 (Rel. 42, Last annotation update)
DE          Serine/chitinase-protein kinase NEK1 (EC 2.7.1.37) (Nima-related
GN          protein kinase 1) (NY-REN-55 antigen).
GE          NEK1 OR KIAA1901.
OS          Homo sapiens (Human).
OC          Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC          Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX          NCBI_TaxID=9606;
RN          [1]
SEQUENCE FROM N.A. (ISOFORM 1).
RC          TISSUE=Brain;
RX          MEDLINE=21456161; PubMed=11572484;
RA          Nagase T., Kikuno R., Ohara O.;
RT          "Prediction of the coding sequences of unidentified human genes. XXI.
RT          The complete sequences of 60 new cDNA clones from brain which code for
RT          large proteins.";
RL          DNA Res. 8:179-187(2001).
RN          [2]
SEQUENCE OF 444-1258 FROM N.A. (ISOFORM 2).
RC          TISSUE=Renal cell carcinoma;
RX          MEDLINE=9943124; PubMed=10508479;
RA          Scallan M.J., Gordan J.D., Williamson B., Stockert E., Bander N.H.,
RA          Jongeneel C.V., Gure A.O., Jager D., Jager E., Knuth A., Chen Y.-T.,
RA          Old L.J.;
RT          "Antigens recognized by autologous antibody in patients with
RT          renal-cell carcinoma.";
RL          Int. J. Cancer 83:456-464(1999).
CC          -!- FUNCTION: PHOSPHORYLATES SRINES AND THREONINES, BUT ALSO APPEARS
CC          TO POSSESS TYROSINE KINASE ACTIVITY. IMPLICATED IN THE CONTROL OF
CC          METOSIS (By similarity).
CC          -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC          -!- SUBCELLULAR LOCATION: Nuclear (Probable).
CC          -!- ALTERNATIVE PRODUCTS:
CC          Event=Alternative splicing; Named isoforms=2;
CC          Name=1;
CC          IsoId=Q96PY6-1; Sequence=Displayed;
CC          Name=2;
CC          IsoId=Q96PY6-2; Sequence=VSP_004870;
CC          Note=No experimental confirmation available;
CC          -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA
CC          subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announcement/
or send an email to licenses@isb-sib.ch).
-----
EMBL; AB067488; BAB67794.1; ALT_INIT.
EMBL; AF155113; AAD42879.1; -.
GENE; HGNC:7744; NEK1.
CC          GK; Q96PY6; -.
CC          MIM; 604588; -.
CC          InterPro; IPR000719; Prot Kinase.
CC          InterPro; IPR008271; Ser Thr pkin AS.
CC          InterPro; IPR002290; Ser_thr_kinase.
CC          InterPro; IPR001245; Tyr_kinase.
CC          Pfam; PF00069; pkinase; 1.
CC          PRINTS; PR00109; TYRKINASE.
CC          ProDom; PD000001; Prot Kinase; 1.
CC          SMART; SM00220; S_TKc; 1.
CC          PROSITE; PS00107; PROTEIN KINASE ATP; 1.

```

DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Nuclear protein; Phosphorylation; Cell cycle; Cell division;
KW Tyrosine-protein kinase; Alternating splicing.
FT DOMAIN 4 258 PROTEIN_KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 128 128 BY SIMILARITY.
FT MOD_RES 162 162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT VARSPLIC 478 521 Missing (in isoform 2).
FT FTIG=VSP 004870.
FT CONFLICT 1232 1232 G -> E (IN REF. 2).
SQ SEQUENCE 1258 AA; 142828 MW; 339C4BFA56612530 CRC64;
Query Match 29.4%; Score 1063.5; DB 1; Length 1258;
Best Local Similarity 33.4%; Pred. No. 8.7e-50;
Matches 279; Conservative 122; Mismatches 232; Indels 203; Gaps 23;
QY 1 MDKVDVIAIGAGFGKAVLAGKSDSKHCIVKEINFEKMPQKEASKKVLLEKMKH 60
DQ 1 MEKVRQKIGSGFGKAILVKSTEDGQYVKEINISRMSSKREESRRVAVLANWKH 60
QY 61 PNIVAFNFSQBNGLRFIMVEYCDGDLMKRINRGVLFSEDFILGHFVQISLGLKH 120
DQ 61 PNIVQVRESFEENGLYIMDYCEGDLFKRINAGKGVLFQEDQLDFVQICLAKH 120
QY 121 DKILHRDIKAIIFLSKNGMTAKLGDGCIARVLANSMELARTCIGTPYLSPEICNKP 180
DQ 121 DKILHRDIKAIIFLSKNGMTAKLGDGCIARVLANSMELARTCIGTPYLSPEICNKP 179
QY 181 YNNKTDIWSLGLVVELCTLKHFFEGNNLQQLVLCIQAHFAPISGFSRSLHSLSOLF 240
DQ 180 YNNKSDIWSLGLVVELCTLKHFFEGNNLQQLVLCIQAHFAPISGFSRSLHSLSOLF 239
QY 241 QVSPDRPSINSILKRPFLNLIPIKLTPEVIOEFSHMLICRAGA---PSRHAG--- 293
DQ 240 KENPRDRPNSVILEKGFIAKIEKFLSPQLAEFCILKTSFGSQIPAKRPSAGONS 299
QY 294 -----KVVQCKIKQVRFQCKPPRSRISVPTKRWAI 325
DQ 300 ISVMPAQKITPAKXVIGPLAYKKGDKLHEKPLQXHK QAHQTPKRVNTGEERKI 358
QY 326 -----LHRNE-----WRPPAGAQA 340
DQ 359 SBEAARKRLTEFEKEKQKQDIISLMKAQKQKQERLERINRAREQGNVLSAGGS 418
QY 341 RSIKM-----IEPKIAAVGHDYDYAQLD-MLRRRAHK-----PSYH 378
DQ 419 GEVKAPFLGGGTIAPSSFSRSGQYEHVHAIFDQMQQRAEDNEAKWKREYGRGLPERG 478
QY 379 PIPQENTGVEDYGOETRH-----GFSQSPQAPAY 407
DQ 479 ILPGVRPGF-PYGAAGHHFPDADDIKTLKELKAVSQANRQKQLAVERAKQVEEF 537
QY 408 LQKFEAQYKLVKVE-----KQLGRPSAEPNPNORQELRS--NGEPRQELPF 456
DQ 538 LQKFEAQYKLVKVE-----KQLGRPSAEPNPNORQELRS--NGEPRQELPF 456
QY 457 RKNEMKEQYVWQLEIRQYHND---MKEIRKKNGREPEENSKISHTKYLKKS----- 508
DQ 592 QEGSEADMRKIESLKAHANARAVALKEQLERKKEAYEREKVVBEHLVAKGVKSSD 651
QY 509 -NLPHQDASEGAPVQ-----DIEKDLKQM-----RLQNTKESKNPQK-----YKAKG 553
DQ 652 VSPPLQGHGTGSPSQQRVSVISVTSALKYGVDSLSLTDRETSEEMQKTNNAISSRE 711
QY 554 VKFEINLDCISDENTLQEEAMDIENETLTEDGKPKFKEYECVKEHGDYTDK-----APE 609
DQ 712 ILRLN-----ENLKAQEDKGNLSDTTEINVHEDAKHEKEKSVSSDRKKWEAGG 764
QY 610 KLHCP-----EAGFSTQVAAVGN-----RRQWDGAPOTLLQMMVAADI 649

Db 765 QLVIPDELTLDTSFSTTERTHTVGEVILKPNQSPRAWKSPSTDVSLKILGEAEL 820
RESULT 3
NEK3 MOUSE
ID NEK3 MOUSE STANDARD; PRT; 511 AA.
AC Q9ROA5; Q9ZOX9.
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase Nek3 (EC 2.7.1.37) (NIMA-related protein kinase 3).
GN NEK3.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=99240743; PubMed=10224116;
RA Tanaka K., Nigg E.A.;
RT "Cloning and characterization of the murine Nek3 protein kinase, a novel member of the NIMA family of putative cell cycle regulators.";
RL J. Biol. Chem. 274:13491-13497(1999).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=99321807; PubMed=10393247;
RA Chen A., Yanai A., Arama E., Kilfin G., Motro B.;
RT "NIMA-related kinases: isolation and characterization of murine nek3 and nek4 cDNAs, and chromosomal localization of nek1, nek2 and nek3.";
RL Gene 234:127-137(1999).
CC CC
CC -!- FUNCTION: Kinase that may play a role in mitotic regulation.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
CC
CC EMBL; AF093416; AAD20986.1; -
CC EMBL; AF099066; AAD16286.1; -
CC HSP; O63450; 1A06.
CC MGD; MGI:1344371; Nek3.
CC InterPro; IPR000719; Prot_kinase.
CC InterPro; IPR008271; Ser_Thr_kinase.
CC InterPro; IPR002290; Ser_Thr_kinase.
CC InterPro; IPR001245; Tyr_kinase.
CC Pfam; PF00069; pkinase; 1.
CC PRINTS; PR00109; TYRKINASE.
CC ProDom; PD000001; Prot_kinase; 1.
CC SMART; SM00220; S_TKc_1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS50011; PROTEIN_KINASE_DOM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Nuclear protein; Phosphorylation; Cell cycle; Cell division.
FT DOMAIN 4 255 PROTEIN_KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT BINDING 33 33 ATP (BY SIMILARITY).
FT ACT_SITE 125 125 BY SIMILARITY.
FT MOD_RES 159 159 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
FT CONFLICT 239 239 N -> K (IN REF. 2).
FT CONFLICT 342 343 MISSING (IN REF. 2).
SQ SEQUENCE 511 AA; 57222 MW; DB6D6C0533C7302F CRC64;
Query Match 22.4%; Score 808.5; DB 1; Length 511;
Best Local Similarity 33.1%; Pred. No. 1.4e-36;
Matches 202; Conservative 99; Mismatches 193; Indels 117; Gaps 16;

QY 1 MDKYDVKAIGOGARGKAVLAKGSDSKHCYKVINFEKMPQIOEKASKEVILEKMKH 60
 Db 1 MONTYVLRVIGGSGFRALLVQESNQTFAMKEIRLLK---SDTQTSRKEAVLLAKMKH 57
 QY 61 PNIVAFPSFQNGRLFIWMEYCDGDLMKRINRQGVLFSDQILGWVFOISGLKXIH 120
 Db 58 PNIVAFKSEFEAGVLYIWMEYCDGDLMKRQKQGNLFPEDTILNWFQICLVGNH 117
 QY 121 DRKILHRDIAQNIPLSKNGVAKLGDGFIARVLNNSMELARTCTGTPYILSPICQNP 180
 Db 118 KRVLRHRDIAQNVFLTHNGV-KLGDGFIARVLNNSMELARTCTGTPYILSPICQNP 176
 QY 181 YNNKTDIWSLGVLYELCTLKHPPGNNLQQLVKICQAHFAPISPGFSRELHSLISOLF 240
 Db 177 YNNKTDIWSLGVLYELCTLKHPPGNNLQQLVKICQAHFAPISPGFSRELHSLISOLF 236
 QY 241 QVSPDRPINSILKRPFLNIPKYLTPTEVTOEFHMLICRAGAPASRHAGKVKQCK 300
 Db 237 KENPSHRPSATLLCRGSLAPLVKLPPIRE-----YGEQILDEIK 280
 QY 301 IQKVRFOGKCPPRSIRISVPIKENAILHRNEWRPPAGACKARSIKMIERPKIAAVCGHYD 360
 Db 281 ISTPKNMK-----QDSNRVGRALGEANSAAMQEEER-GRKCSHTE- 320
 QY 361 YYAQLDMLRRRAHKPSYHPIPOENTGVEDYQOETHGPPSPQWPAEYLQKFEAQOYK 420
 Db 321 -----LESTGTTAGNALGAARGNPESGNRQHGSHTS-PASPHRPRWERH----- 366
 QY 421 VEKQLGLSPSSAEPYNNQRLSRNGEPRQELPFRKNMEQOYKWLBEIISQOYHND 480
 Db 367 -----GPSSNVLEAKASILTSS-----FTAEDRGGSVIKYEENARQWRE 409
 QY 481 MKEIRKMGREPEENS KISHKTYLVKXGNL-----PVHQDASGEAPVQDIEXDLQMR 534
 Db 410 PPEALLSLMKADLQ--AFQTYIYRFGAEGFLKGLPSEDTAS-----DSVDGLDSVM 462
 QY 535 LQNTKESKNPEQYKAKGVKFEINLDRKICSDENILQEEEMDPTNETLTFEDGMKFEY 594
 Db 463 L-----DPE-----RFEPLDDE--SDTQFEDNPNPWSE----- 491
 QY 595 ECVKEHGYD 605
 Db 492 --LKXHVGYD 500
 RESULT 4
 ID NEK3 HUMAN STANDARD; PRT; 506 AA.
 AC P51956; QBMUN5;
 DT 01-OCT-1996 (Rel. 34, Created)
 DT 28-FEB-2003 (Rel. 41, Last sequence update)
 DT 15-MAR-2004 (Rel. 43, Last annotation update)
 DE Serine/threonine-protein kinase Nek3 (EC 2.7.1.37) (Nima-related
 protein kinase 3) (HSPK 36).
 GN NEK3.
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 OX NCBI_taxid=9606;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins B.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altshul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Shat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heide F.,
 RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hong F.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Ustin T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaby S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,

RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Whiting J., Heiton E., Kettelman M., Maman A., Rodrigues S., Sanchez A.,
 RA Blakesley R.W., Tuchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalios D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
 RT "Generation and initial analysis of more than 15,000 full-length
 human and mouse cDNA sequences.";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP SEQUENCE OF 48-506 FROM N.A.
 RC TISSUE=Placenta;
 RX MEDLINE=94368699; PubMed=7522034;
 RA Schults S.J., Fry A.M., Suetterlin C., Ried T., Nigg E.A.;
 RT "Cell cycle-dependent expression of Nek2, a novel human protein
 kinase related to the NIMA mitotic regulator of Aspergillus
 nidulans.";
 RL Cell Growth Differ. 5:625-635 (1994).
 RN [3]
 RP SEQUENCE OF 76-189 FROM N.A.
 RX MEDLINE=94100173; PubMed=8274451;
 RA Schults S.J., Nigg E.A.;
 RT "Identification of 21 novel human protein kinases, including 3
 members of a family related to the cell cycle regulator nima of
 Aspergillus nidulans.";
 RL Cell Growth Differ. 4:821-830 (1993).
 CC -!- FUNCTION: Kinase that may play a role in mitotic regulation.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA
 subfamily.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 between the Swiss Institute of Bioinformatics and the EMBL outstation -
 the European Bioinformatics Institute. There are no restrictions on its
 use by non-profit institutions as long as its content is in no way
 modified and this statement is not removed. Usage by and for commercial
 entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; BC019916; AAH19916.2; -;
 DR EMBL; Z29067; CA82310.1; -;
 DR EMBL; Z25434; CA80921.1; -;
 DR FIR; I38224; I38224.
 DR HSSP; Q00534; IBI7.
 DR Genew; HGNC:7746; NEK3.
 DR GK; P51956; -;
 DR MIM; 604044; -;
 DR GO; GO:0005634; C:nucleus; NAS.
 DR GO; GO:0005824; F:ATP binding; NAS.
 DR GO; GO:0004674; F:protein serine/threonine kinase activity; NAS.
 DR GO; GO:0007049; P:cell cycle; NAS.
 DR GO; GO:0007067; P:mitosis; NAS.
 DR GO; GO:0006468; P:protein amino acid phosphorylation; NAS.
 DR InterPro; IPR000719; Prot kinase.
 DR InterPro; IPR008271; Ser Thr pkin AS.
 DR InterPro; IPR002290; Ser_thr_kinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot kinase; 1.
 DR SMART; SM00220; S_TK; 1.
 DR PROSITE; PS00107; PROTEIN KINASE ATP; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE-ST; 1.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
 FT Nuclear protein; Phosphorylation; Cell cycle; Cell division.
 FT DOMAIN 4 257 PROTEIN KINASE.
 FT NP BIND 10 18 ATP (BY SIMILARITY).
 FT BINDING 33 33 ATP (BY SIMILARITY).
 FT ACT SITE 127 127 BY SIMILARITY.
 FT MOD_RES 161 161 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).

```
FT CONFLICT 54 54 L -> S (IN REF. 2).
FT CONFLICT 77 78 IV -> LV (IN REF. 3).
FT CONFLICT 187 189 SIG -> RSV (IN REF. 3).
SQ SEQUENCE 506 AA; 57704 MW; 4437EB441A44777 CRC64;

Query Match
Best Local Similarity 34.1%; Pred. No. 1.5e-36;
Matches 196; Conservative 97; Mismatches 180; Indels 102; Gaps 16;

QY 1 MDYDVIKAGGAGKAVLAKGSDSKHCIVKEINFEKMPQIOEAKSKVEVILLKMKH 60
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 61 PNIVAFNSFQNGRLFIWVEYCDGDLMKRINRQGVLFSDQILGHFVQI-SLGLKH 120
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 121 DKILHRDIKAQNIPLSKNGWAKLGDGFIARVLNNSMELATCTGTGTPYVSPET 180
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 120 KCVLHRDIKSNIEFLTQNGKV-KLGDGFSARLSNPMAFACVYVGTTPYVPEI 178
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 181 YNNKTDIWSLGVLYELCTLKHPPFGNNLQQLVKIQAHFAPISPGFSRELHSL 240
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 179 YNNKSDIWSLGVLYELCTLKHPPFQANSWNILKVCQCISPLSHYSYELQFLV 238
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 241 QVSPDRPINSILKRPFLNIPKYLTPVTOEFSHMLICRAGAPASRHAGKYVQK 300
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 239 KNPGRPSATLLSGRIVARLVQKLPPEITME-----YGEVLEBIK 282
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 301 IQKVRPQGCPRPSRISVPIKENALHRNEWPPAGAKARSIKMIERPKAACVGH 360
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 283 NKNHTPRKTNPSRIRIALGNEASTVQEEQDRKGS----- 319
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 361 YYAQLDMLRRRAKPSYHP IPOENTGVEDYQGTGRHGPS-----QWPAEYL 416
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 417 YKLVEKQLGLPPSAEPNPNQRLSN--GEPR-FQELPFRKNEMKEQYWKQLEI 473
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 364 -----EKNV---PNTALTALENASILTSLTAEDDRGGSVIKSKNTTRKQ-- 411
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 474 RQCYHNDMKEIRKMGREPEENSKI-SHKTLYV-----KKNLPVHQDASEG-- 520
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 412 PDLTLNLK-----NADLSLAFQTYTYRPGSEGLKGLPSETEASDSVDGGH 461
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 521 APVQD---IEKDLKQMLQNTKESKNPQKYAKK 552
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 462 SVILDPERLEPGLDEEDTDFEEDNDPNWVSELKK 496
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

RESULT 5
NEK4_MOUSE
ID NEK4_MOUSE STANDARD; PRT; 792 AA.
AC Q921J2; O35673; Q9RLJ1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase Nek4 (EC 2.7.1.37) (NIMA-related protein kinase 4) (Serine/threonine protein kinase 2).
GN NEK4 OR STK2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]_
RP SEQUENCE FROM N.A. (ISOFORMS 1 AND 2).
RC STRAIN=BALB/c; TISSUE=Spleen;
RX MEDLINE=20001940; PubMed=10529384;
RA Hayashi K., Igarashi H., Ogawa M., Sakaguchi N.;
RT "Activity and substrate specificity of the murine STK2
RT serine/threonine kinase that is structurally related to the mitotic
RT regulator protein NIMA of Aspergillus nidulans.";
RT Biochem. Biophys. Res. Commun. 264:449-456(1999).
RL
```

```

[2]
SEQUENCE FROM N.A. (ISOFORM 1).
MEDLINE=9321807; PubMed=10393247;
Chen A., Yanai A., Arama E., Kilfin G., Motro B.;
"NIMA-related kinases: isolation and characterization of murine nek3,
and nek4 cDNAs, and chromosomal localization of nek1, nek2 and nek3.";
Gene 234:127-137(1999).
-!- FUNCTION: Seems to act exclusively upon threonine residues.
-!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
-!- SUBCELLULAR LOCATION: Nuclear (Probable).
-!- ALTERNATIVE PRODUCTS:
Event=Alternative splicing; Named isoforms=2;
Name=1; Synonyms=mSTK2L;
IsoId=Q921J2-1; Sequence=Displayed;
Name=2; Synonyms=mSTK2S;
IsoId=Q921J2-2; Sequence=VSP_007001;
-!- TISSUE SPECIFICITY: Expressed ubiquitously among various organs
and is upregulated in the testis.
-!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA
subfamily.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (See http://www.isb-sib.ch/announce/
or send an email to license@isb-sib.ch).
-----
EMBL; AJ232071; CAAL1072.1; -.
EMBL; Y05234; CAA70436.1; -.
EMBL; AF099067; AAD16287.1; -.
PIR; JC7122; JC7122.
DR HSSP; Q63450; IAC06.
MGD; MGI:1344404; Nek4.
InterPro; IPR000719; Prot_kinase.
InterPro; IPR008271; Ser_Thr_pkin_AS.
InterPro; IPR002290; Ser_Thr_pkinase.
InterPro; IPR001245; Tyr_pkinase.
Pfam; PF00069; pkinase; 1.
PRINTS; PR00109; TYRKINASE.
ProDom; PD000001; Prot_kinase; 1.
SMART; SM00220; S_TKc_1.
PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
Nuclear protein; Phosphorylation; Alternative splicing.
DOMAIN 6 261
NP_BIND 12 20 ATP (BY SIMILARITY).
BINDING 35 35 ATP (BY SIMILARITY).
ACT_SITE 131 131 BY SIMILARITY.
MOD_RES 165 165 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
VARSPIC 456 503 Missing (in isoform 2).
/FTid=VSP_007001.
G -> R (IN REF. 2).
CONFLICT 499 499 88994 MW; CF9187311C807ALD CRC64;
SQ SEQUENCE 792 AA;

Query Match 18.4%; Score 665.5; DB 1; Length 792;
Best Local Similarity 25.3%; Pred. No. 1.1e-28;
Matches 189; Conservative 137; Mismatches 213; Indels 209; Gaps 23;

QY 4 YDVYKAIGGAGKAVLAKGSDSKHCIVKEINFEKMPQIOEAKSKVEVILLKMKHPNI 63
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 6 YCVMVVRGSGYGEVILVKKHRDGGKQYVIKKILNRNASSERRAARAEQALLQLKHPNI 65
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 64 VAFNSFQ-ENGRLFIWVEYCDGDLMKRINRQGVLFSDQILGHFVQI-SLGLKH 122
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 66 VTYKESMEGGDGLIYVWGFCGGDLRYLKEQGLLPESQVVEVFWQIAMAQLVLEH 125
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 123 KILHRDIKAQNIPLSKNGWAKLGDGFIARVLNNSMELATCTGTGTPYVSPET 182
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:
QY 126 HILHRDQKTNVFLTRNII-KVGDGLIARVLEHMGDMASTLIGTFYVWSPFLSKNPTN 184
Db |||:||||:||||:||||:||||:||||:||||:||||:||||:||||:||||:

```



```
QY 389 DYQG-----ETHGFSPS-----QW-----PAEYLQKRFQAQYK-LK----- 420
Db 402 SISQVEEEMLDNTRSSAQENLIPWMSDDIVTGKNEFVKPLOPLIKEQKPKQSLALS 461
QY 421 -----VEKQLGLRP-- 429
Db 462 PKLECSGTILAHNURLGSSDSPASASRVAGITGVCHHAQDQVAGECIIIEKQRIHPDL 521
QY 430 ----SAPPNYN-QQELRSNGEPR-----FQELPFR-----KN 459
Db 522 QPHNSGSBPSLSRQKRRQREQTHRGKQVRRDLFAFQESPFRFLPSHPIVGKVDVTS 581
QY 460 ENKEOYKQ-----LEIRQYHNDKE-----IRKMGREPEENSKIHKYLVYKKS 508
Db 582 TOKENRQVVTGVSSSRSSEMSSKDRPLSARERRRLKQSGEMSSSPS-----VRKA 638
QY 509 NLPVHQDASGEAPVQD-----IEKDLKQML-----QNTRESKNPEQY 548
Db 639 SLSV---AGPGKQEQDPLPARRLSSDCSVTQERKQIHCLSEDFLSSSTSTSDKSDGY 695
QY 549 KAKGVKFEINLDKISDENIL--QEEEMDIP--NETLTEDGKFKFKEVCVEHGDYT 604
Db 696 GSGKQCTNEINALVOLMTQTLLKDSKESCEDVPVANPVSEFKLHRKYRDTLIL--HGKVA 753
QY 605 DXAFKELHCP 615
Db 754 EEA-EHIFKE 763

RESULT 7
NEK8 HUMAN
ID NEK8 HUMAN STANDARD; PRT; 692 AA.
AC Q86566; Q8NDH1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DE Serine/threonine-protein kinase Nek8 (EC 2.7.1.37) (Nima-related
protein kinase 8) (NIMA-related kinase 12a).
GN NEK8 OR NEK12A.
OS Homo sapiens (Human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxId=9606;
RN [1]
RP SEQUENCE FROM N.A.
RA Huang C.Q., Wu S.L., Yu L.;
RL Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Nedwiedek M.N., Roig J., Lenz G., Avruch J.;
RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE OF 445-692 FROM N.A.
RC Tissue-Testis;
RA Ottenwelder B., Obermaier B., Mewes H.-W., Gassenhuber J.,
RA Wiemann S.;
RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
CC -!- FUNCTION: Required for renal tubular integrity (By similarity).
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBCELLULAR LOCATION: Predominantly cytoplasmic (By similarity).
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA
subfamily.
CC -!- SIMILARITY: Contains 5 RCC1 repeats.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; AY242354; AAC88243.1; -.
```

```
DR EMBL; AY267371; AAP04006.1; -.
DR EMBL; AL833909; CAD38765.1; -.
DR Genew; HGNC:13387; NEK8.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR000408; Reg_Chr_condens.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR InterPro; IPR001245; Tyr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR Pfam; PF00415; RCC1; 1.
DR PRINTS; PR00633; RCNDNSATN.
DR PRINTS; PR00109; TYRKINASE.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR SMART; SM00219; TYFKC; 1.
DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00625; RCC1_1_FALSE_NEG.
DR PROSITE; PS00636; RCC1_2_FALSE_NEG.
DR PROSITE; PS00012; RCC1_3_5.
KW Transferase; Serine/threonine-protein kinase; ATP-binding;
KW Phosphorylation; Repeat.
FT DOMAIN 4 258 PROTEIN_KINASE.
FT NP_BIND 10 18 ATP (BY SIMILARITY).
FT REPEAT 312 350 RCC1 1.
FT REPEAT 410 461 RCC1 2.
FT REPEAT 462 513 RCC1 3.
FT REPEAT 580 631 RCC1 4.
FT REPEAT 632 684 RCC1 5.
FT ACT_SITE 128 128 BY SIMILARITY.
FT BINDING 33 33 ATP (BY SIMILARITY).
FT MOD_RES 162 162 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 692 AA; 74806 MW; 9E09820DFB3D5CA1 CRC64;

Query Match 17.5%; Score 634; DB 1; Length 692;
Best Local Similarity 38.9%; Pred. No. 4.5e-27;
Matches 147; Conservative 63; Mismatches 128; Indels 40; Gaps 8;

QY 1 MDXYDVTKATGCGAFGKAVLAKGKSDSKHCVKEINFERKMPQKEASKEVILLKMKH 60
Db 1 MEKIERIRVVRGAGFIVHLCLRKADQKLVILKQIPVEQTKERQAQNECOVLKLNH 60
QY 61 PNIVAFNSFQENGRLFIWMEYCDGGLDMKRINRQGVLFSEDIILGWVQVSLGKHHT 120
Db 61 PNVEYVENFLEDKALMIAMEYAPGGTAEFIQKRCNSLLEETILHFFVQILLALHHV 120
QY 121 DRKILHRDIAKONIFLSKNGWAKLGDGFIARVNNMSMELARTCIGTPYLSPEIQNK 180
Db 121 THLILHRDLKTONILLDKHRMVVKIGDFGISKIL-SSSKAYTVVGTGTCYISPELCEGP 179
QY 181 YNNKTDIWSLGCVLVEICTLKHFFEGNNLQQLVKICQAHFAPISPGFSRELHSLISQL 240
Db 180 YNQKSDIWLGCVLVELASLKRAFEANLPAVLKIMSGTFAPISDRYSPQLVLSLL 239
QY 241 QVSPRRPSPINSILKRPFLLENLPKYLTFEVTQEEFESHMLICRAGAPASHACKVQK-- 298
Db 240 SLEPACRPPLSHIMAOPLCIRALLNLHTD-----VGSVRVRRAEKSVAPSN 285
QY 299 --CKIQKRVQG--KCPPRSISVPITK-----RNAILHRNWRPPAG-AQKA 340
Db 286 TGSRTTSVRCRGIPRGFVRPAIPPLSSVYVWGGLGTPLRLPMLNTEVVQVAGRTQKA 345
QY 341 ---RSIKMI--ERPRTAA 353
Db 346 GVTRSGRLILWEAPPLGA 363

RESULT 8
NEK8 MOUSE
ID NEK8 MOUSE STANDARD; PRT; 698 AA.
AC Q91ZR4; Q9D685;
DT 10-OCT-2003 (Rel. 42, Created)
```

10-OCT-2003 (Rel. 42, Last sequence update)
 10-OCT-2003 (Rel. 42, Last annotation update)
 Serine/threonine-protein kinase Nek8 (EC 2.7.1.37) (NIMA-related protein kinase 8).
 NEK8.
 Mus musculus (Mouse).
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NCBI_TaxID=10090;
 [1]_TaxID=10090;
 SEQUENCE FROM N.A. (ISOFORM 1), MUTAGENESIS OF LYS-33, AND VARIANT ARKPD VAL-448.
 STRAIN=C57BL/6J; PubMed=12421721;
 MEDLINE=22308894; PubMed=12421721;
 Liu S., Lu W., Obara T., Kuida S., Lehoczy J., Dewar K.,
 Drummond I.A., Beier D.R.;
 "A defect in a novel Nek-family kinase causes cystic kidney disease in the mouse and in zebrafish."
 Development 129:5839-5846 (2002).
 [2]
 SEQUENCE FROM N.A. (ISOFORM 2).
 STRAIN=C57BL/6J; TIGR=SUE-Skin;
 MEDLINE=21085660; PubMed=11217851;
 Kawai J., Shinagawa A., Shibata K., Yoshino M., Itoh M., Ishii Y.,
 Arakawa T., Hara A., Fukunishi Y., Konno H., Adachi J., Fukuda S.,
 Aizawa K., Izawa M., Nishi K., Kiyosawa H., Kondo S., Yamataka I.,
 Saito T., Okazaki Y., Gojibori T., Bono H., Kasukawa T., Saito R.,
 Kadota K., Matsuda H.A., Ashburner M., Batalov S., Casavant T.,
 Fleischmann W., Gaasterland T., Gissi C., King B., Kochiwa H.,
 Kuehl P., Lewis S., Matsuo Y., Nikaide I., Pesole G., Quackenbush J.,
 Schriml L.M., Stauber F., Suzuki R., Tonita M., Wagner L., Washio T.,
 Sakai K., Okido T., Furuno M., Aono H., Baldarelli R., Barsh G.,
 Blake J., Boffelli D., Bojunga N., Carninci P., Fujita M., Gariboldi M.,
 Brownstein M.J., Bult C., Fleischer M., Hume D.A., Kaniya M., Lee N.H.,
 Gustincich S., Hill D., Hofmann M., Hume D.A., Kaniya M., Lee N.H.,
 Lyons P., Marchionni L., Mashima J., Mazzarelli J., Mombaerts P.,
 Nordone P., Ring B., Ringwald M., Rodriguez I., Sakamoto N.,
 Sasaki H., Sato K., Schoenbach C., Seva T., Shibata Y., Storch K.-F.,
 Suzuki H., Toyooka K., Wang K.H., Weitz C., Whitaker C., Wilming L.,
 Wyshaw-Boris A., Yoshida K., Hasegawa Y., Kawaji H., Kotsuki S.,
 Hayashizaki Y.;
 "Functional annotation of a full-length mouse cDNA collection."
 Nature 409:685-690 (2001).
 RT Nucleotide annotation of a full-length mouse cDNA collection."
 CC -!- FUNCTION: Required for renal tubular integrity. May regulate local cytoskeletal structure in kidney tubule epithelial cells.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- SUBCELLULAR LOCATION: Predominantly cytoplasmic.
 CC -!- ALTERNATIVE PRODUCTS:
 Event=Alternative splicing; Named isoforms=2;
 Name=1;
 IsoId=Q912R4-1; Sequence=Displayed;
 Name=2;
 IsoId=Q912R4-2; Sequence=VSP_007959, VSP_007960;
 Note=Due to intron retention. No experimental confirmation available;
 CC -!- TISSUE SPECIFICITY: Kidney, liver, and testis.
 CC -!- DISEASE: Defects in NEK8 are the cause of autosomal recessive juvenile polycystic kidney disease (ARPKD).
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA subfamily.
 CC -!- SIMILARITY: Contains 5 RCC1 repeats.
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL Outstation - European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; AF407579; AAL09675.1; -
 DR EMBL; AK014546; BAB29424.1; -
 DR MGD; MGI:1890646; Nek8.

DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR000408; Reg Chr condens.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR PRINTS; PR00415; RCC1; 3.
 DR PRINTS; PR00633; RCCNDNGATION.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKc_1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; 1.
 DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
 DR PROSITE; PS00625; RCC1_1; FALSE_NEG.
 DR PROSITE; PS00626; RCC1_2; FALSE_NEG.
 DR PROSITE; PS00012; RCC1_3; 5.
 DR Transferase; Serine/threonine-protein kinase; ATP-binding; Phosphorylation; Repeat; Alternative splicing; Disease mutation.
 KW DOMAIN 4 258
 FT NP_BIND 10 18
 FT REPEAT 416 467
 FT REPEAT 468 519
 FT REPEAT 520 595
 FT REPEAT 586 637
 FT REPEAT 638 690
 FT ACT_SITE 128 128
 FT BINDING 33 33
 FT MOD_RES 162 162
 FT VARSPPLIC 277 291
 FT
 FT VARSPLIC 292 698
 FT VARIANT 448 448
 FT MUTAGEN 33 33
 FT CONFLICT 15 15 F -> L (IN REF. 2).
 FT CONFLICT 246 246 R -> G (IN REF. 2).
 FT SEQUENCE 698 AA; 75264 MW; 686B29A8CF180E94 CRC64;
 Query Match 17.5%; Score 633; DB 1; Length 698;
 Best Local Similarity 38.0%; Pred. No. 5.2e-27;
 Matches 147; Conservative 65; Mismatches 123; Indels 52; Gaps 9;
 QY 1 MDKYDVIRKAGGAGKAVLAKGSDSKHCVTKKEINFEKMPQOEKASKVILEKMKH 60
 Db 1 MEKYEIRVVGAGFGIVHLCKRADQKLVILKQIPVEQMTKEERQAAQNCQVLKLNH 60
 QY 61 PNIVAFNFSQNGRLFTVMEYCDGDLKTRINRQGVLFSEDDQILGWFIQISLKHIIH 120
 Db 61 PNIVYEYENFLEDKALMIAMEYAPGGTLAEFTQKCNLSLEETILHFFVQILLALHVV 120
 QY 121 DRKILHRDIKQNIPLSKNGWAKLGDGFIARVLNNSNELATCTGTPYLSPELCKNP 180
 Db 121 THLILHRDLKTONILLDKHRMVKVIGDFGISKIL-SSKSKAYTVVGTTCYISPELCEGP 179
 QY 181 YNNKTDIVSLGCVLYELCTLKHPPFEGNNLQQLVLIKQAHFAPISPGFSRHLHSLSOLF 240
 Db 180 YNQSDDIHALGCVLYELASLKEAFEAANLPALVLMKSGTFAPISDRYSPELRQLVLSLL 239
 QY 241 QVSPDRDRSINSILKRP-LENI-----IPKYLTPETVIEEFSHMLICRAG 285
 Db 240 SLEPQRPLSHIMAPLCIRALINHTDVGSRMRRAEKLTP-----G 284
 QY 286 AFASHAGKGVQCKIQVFOGKCPPRSISVPIK-----RNAILHRNEWR 332
 Db 285 PPIA--SGTSGRATSACRGVPGFVPAFPPLPSSVYANGGLSSPLRLPMLNTEVVQ 342
 QY 333 PPAG-AQKA---RSIKMI--ERPRTAA 353
 Db 343 VAAGRTQKAGVTRSGRLILWEAPPLGA 369

Query Match	17.2%	Score 623;	DB 1;	Length 697;
Best local Similarity	47.0%;	Pred. No. 1.8e-26;		
Matches 124;	Conservative 46;	Mismatches 88;	Indels 6;	Gaps 2;

QY	1	MDXYDVILKIGCGAGFKAYLAKGSKDSKCHVKEIKNEFKPIOEKEASKEVILLERWKH	60
DB	1	MEYVEKTVVGRGAFGVHLCRRTTSALVILKEIPVEQWTRDERLRAQNECQVKULSH	60
QY	61	PNIVAFNSFQENGRULFIWVEYCDGGLMKRINRQGVLPSEDOILLGWFVQISLGLKHH	120
DB	61	PNIIETYYENFLEDKALMIAMEYAPGGTLDIYIOKRCNSLDEDTILHSPVQIILLALYHVH	120
QY	121	DRKILHRIDKAOINFSLSGMWAKLGDGFIARVYLNNSMELARTICIGTFPYVLSPEICQKP	180
DB	121	NKLIHRDLKTQNLIDKQMTIVKIGDFGISKIL-VSKSKAYIVVGTPCVISPELCEGKP	179
QY	181	YNNKTDIWSLGVLYBELCTLKHPFEGNNLQQLVLYKICQAHFAPISPGFSBELHSLISOLF	240
DB	180	YNOKSDIWAIGCVLYELASLKRAFEAAANLPALVYKIMSGTFAPISDRYSPELRQLILNML	233
QY	241	QVSPDRDRPSINSILK-----RPPL 259	
DB	240	NLDPSKRPQLNEIMAHACIRPLL 263	

RESULT 10				
NEK2 HUMAN				
ID	NEK2 HUMAN	STANDARD;	PRT;	445 AA.
AC	P51955;			
DT	01-OCT-1996 (Rel. 34, Created)			
DT	01-OCT-1996 (Rel. 34, Last sequence update)			
DT	15-MAR-2004 (Rel. 43, Last annotation update)			
DE	Serine/threonine-protein kinase Nek2 (EC 2.7.1.37) (NimaA-related			
DE	protein kinase 2) (NimaA-like protein kinase 1) (HSPK 21).			
GN	NEK2 OR NLK1			
OS	Homo sapiens (Human)			
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;			
OC	Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.			
OX	NCBI_TaxId=9606;			
OX	[1]			
RC	SEQUENCE FROM N.A.			
RC	TISSUE=Nasopharynx, Placenta, and T-cell;			
RA	MEDLINE=94368699; PubMed=7522034;			
RA	Schultz S.J., Fry A.M., Suetterlin C., Ried T., Nigg E.A.;			
RT	"Cell cycle-dependent expression of Nek2, a novel human protein			
RT	kinase related to the NIMA mitotic regulator of Aspergillus			
RT	nidulans."			
RL	Cell Growth Differ. 5:625-635(1994).			
RN	[2]			
RA	SEQUENCE FROM N.A.			
RA	Lu K.P., Hunter T.;			
RT	"Molecular cloning and expression of NLK1, a human NIMA-like kinase.";			
RT	Submitted (JUL-1994) to the EMBL/GenBank/DBJ Databases.			
RN	[3]			
RC	SEQUENCE FROM N.A.			
RC	TISSUE=Skin;			
RA	MEDLINE=22388257; PubMed=12477932;			
RA	Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,			
RA	Klausner R.D., Collins F.S., Wagner L., Shermen C.M., Schuler G.D.,			
RA	Altschul S.F., Zedberg H., Suetow K.H., Schaefer C.F., Bhat N.K.,			
RA	Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,			
RA	Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,			
RA	Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,			
RA	Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,			
RA	Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,			
RA	Bosak S.A., McEwan P.J., McKernan X.J., Malek J.A., Gunaratne P.H.,			
RA	Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,			
RA	Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,			
RA	Fahney J., Helton E., Kettelman M., Madan A., Rodrigues S., Sanchez A.,			
RA	Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,			
RA	Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,			
RA	Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.			

RA Heilig R., Eckenberg R., Petit J.-L., Fonknechten N., Da Silva C.,
RA Cattolico L., Levy M., Barbe V., De Berardinis V., Ureta-Vidal A.,
RA Pelletier E., Vico V., Anthouard V., Rowen L., Madan A., Qin S.,
RA Sun H., Du H., Pepin K., Antigenave F., Robert C., Cruaud C.,
RA Bruels T., Jallou O., Friedlander L., Samson G., Brottier P.,
RA Cure S., Segreus B., Aniere F., Samain S., Crespeau H., Abbasi N.,
RA Alich N., Boscus D., Dickhoff R., Dots M., Dubois I., Friedman C.,
RA Gouvenoux M., James R., Madan A., Mairey-Estrada B., Mangerot S.,
RA Martins N., Menard M., Oztas S., Ratcliffe A., Shaffer T., Trask B.,
RA Vacherie B., Bellemere C., Belser C., Besnard-Gonnet M.,
RA Bartol-Mavel D., Boutard M., Bries-Silla S., Combette S.,
RA Dufosse-Laurent V., Ferron C., Lechaplais C., Lousse S.,
RA Magdelenat G., Pateau E., Petit E., Sirvain-Trukiewicz P., Trybou A.,
RA Vega-Czarny N., Bataille E., Bluet E., Bordelais I., Dubois M.,
RA Dumont C., Guerin T., Haffray S., Hamadi R., Muanga J., Pellouin V.,
RA Robert D., Wunderle B., Gauguier G., Roy A., Sainte-Marthe L.,
RA Verdier J., Verdier-Discala C., Hallier L.W., Fulton L., McPherson J.,
RA Matuda F., Wilson R., Scarcelli C., Gypay G., Wincker P., Saurin W.,
RA Quetier F., Waterston R., Hood L., Weissbach J.,
RT "The DNA sequence and analysis of human chromosome 14,";
RL Nature 421:601-607(2003).
RN [5]
RP SEQUENCE OF 655-979 FROM N.A.
RC TISSUE=Brain;
RX MEDLINE=22388257; PubMed=12477932;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diachenko L., Marusha K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullighy S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield V.S.N., Krzywinski M.J., Skalska U., Smalios D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences,";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: Plectotropic regulator of mitotic progression,
CC participating in the control of spindle dynamics and chromosome
CC separation. Phosphorylates different histones, myelin basic
CC protein, beta-casein, and BICD2. Phosphorylates histone H3 on
CC serine and threonine residues and beta-casein on serine residues.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphorylated
CC autophosphorylation. Activated during mitosis by intramolecular
CC by manganese >> magnesium ions. Sensitive to increasing
CC concentration of detergents. It is not cell-cycle regulated but
CC activity is higher in G0-arrested cells.
CC -!- SUBUNIT: Homodimer. Binds to Ran GTPase. Has a greater affinity
CC for Ran-GDP over Ran-GTP. Interacts with NEX6, NEX7 and BICD2.
CC -!- SUBCELLULAR LOCATION: Cytoplasmic.
CC -!- TISSUE SPECIFICITY: Most abundant in heart, liver, kidney and
CC testis. Also expressed in smooth muscle cells and fibroblasts.
CC -!- DEVELOPMENTAL STAGE: Expression varied mildly across the cell
CC cycle, with highest expression observed in G1 and stationary-phase
CC cells.
CC -!- DOMAIN: Dimerizes through its coiled-coil domain.
CC -!- PTM: Autophosphorylated on serine and threonine residues.
CC -!- PHOSPHORYLATION: Phosphorylated by CDC2 in vitro.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA
CC subfamily.
CC -!- SIMILARITY: Contains 6 RCC1 repeats.
CC -!- CAUTION: Ref.4 (AAD31936) sequence differs from that shown due to
CC erroneous gene model prediction.

This SWISS-PROT entry is copyright. It is produced through a collaboration
between the Swiss Institute of Bioinformatics and the EMBL outstation -
the European Bioinformatics Institute. There are no restrictions on its
use by non-profit institutions as long as its content is in no way
modified and this statement is not removed. Usage by and for commercial
entities requires a license agreement (see <http://www.isb-sib.ch/announce/>
or send an email to license@isb-sib.ch).

EMBL; AY048580; AAL05428.1; -
DR EMBL; AY080896; AAL87410.1; -
DR EMBL; AB082526; BAC02704.1; ALT_INIT.
DR EMBL; AC007055; AAD31936.1; ALT_SEQ.
DR EMBL; AC007055; AAD31938.1; -
DR EMBL; AC007055; AAD31939.1; -
DR EMBL; AC007055; AAD31940.1; -
DR EMBL; BC009336; AAH09336.1; ALT_INIT.
DR Genew; HGNC:18591; NEX9.
DR InterPro; IPR000719; Prot kinase.
DR InterPro; IPR000408; Reg Chr condens.
DR InterPro; IPR008271; Ser_thr_kinase.
DR InterPro; IPR002290; Ser_thr_kinase.
DR InterPro; IPR001245; Tyr_kinase.
DR Pfam; PF00069; kinase; 1.
DR Pfam; PF00415; RCC1; 2.
DR PRINTS; PR00109; TYRKINASE.
DR PRODOM; PD000001; Prot kinase; 1.
DR SMART; SM00220; S_TKc; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
DR PROSITE; PS00011; PROTEIN_KINASE_DOM; 1.
DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.
DR PROSITE; PS00625; RCC1; 1; FALSE_NEG.
DR PROSITE; PS00626; RCC1; 2; FALSE_NEG.
DR PROSITE; PS00012; RCC1; 3; 6.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Phosphorylation; Cell cycle; Cell division; Repeat; Coiled coil.
FT DOMAIN 52 308
FT NP_BIND 58 66
FT REPEAT 388 444
FT REPEAT 445 498
FT REPEAT 499 550
FT REPEAT 551 615
FT REPEAT 616 668
FT REPEAT 669 726
FT DOMAIN 732 891
FT DOMAIN 765 888
FT DOMAIN 892 939
FT DOMAIN 752 760
FT ACT_SITE 176 176
FT BINDING 81 81
FT MOD_RES 210 210
FT MUTAGEN 81 81
FT MUTAGEN 210 210
FT MUTAGEN 214 214
FT CONFLICT 351 351
FT CONFLICT 429 429
FT CONFLICT 967 967
SQ SEQUENCE 979 AA; 107149 MW; FF2486CCS99322CE CRC64;
Query Match 14.8%; Score 535.5; DB 1; Length 979;
Best Local Similarity 40.7%; Pred. No. 1.3e-21;
Matches 105; Conservative 55; Mismatches 95; Indels 3; Gaps 2;

QY 4 YDVIKAGQAFQKAYLAKGSKSCVKEINFEKMPQKEASKEVLEKMGHNI 63
DB 52 YIPRVLGRGAFGEATLYRTEDDSLVWKEVDLTLSKERDALNEIVALLQHDNI 111
QY 64 VAFENSPQENGRFLIVMEYCDGDLMKRNQSGVLFSDDIILGWVQLSLGLKHIDRK 123
DB 112 IAYINFMDMNTLLILEVCNGNGLYDKILRQDKLFEEMVWVYLFQIVSAVSCHKAG 171
QY 124 ILHRDIKAQNIFLSKNGWAKLGDGFIARVLNNSMELARTCTIGTPYLSPEICQNKPEYN 183

```
Db 172 ILHRDIKILNPLTKANLI-KLGDYGLAKKLNSEYVAETLVGTPYMPSPCLCGVKNF 230
Qy 184 KTDIWSLCVAYELCTLKHPPFEGNNLOQLVLKICQAHFA--PISPGFSRELHLSLSQLFQ 241
Db 231 KSDIWAAGCVIFELLTLKRTTDTATPLNLCVKIVQIRAMEVDSSQYSLELIQWVHSCLD 290
Qy 242 VSPDRPSPINSILKRPFL 259
Db 291 QDPEQRPTADELLDRPLL 308

RESULT 12
NEK2_MOUSE
ID NEK2_MOUSE STANDARD; PRT; 443 AA.
AC Q35942; O35959;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase Nek2 (EC 2.7.1.37) (Nima-related
DE protein kinase 2).
GN NEK2.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
RN [1]
RP SEQUENCE FROM N.A., SUBCELLULAR LOCATION, AND TISSUE SPECIFICITY.
RC STRAIN=Swiss Webster; TISSUE=Testis;
RX MEDLINE=97330684; PubMed=9187143;
RA Rhee K., Wolgemuth D.J.;
RT "The NIMA-related kinase 2, Nek2, is expressed in specific stages of
RT the meiotic cell cycle and associates with meiotic chromosomes.";
RL Development 124:2167-2177(1997).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=98243037; PubMed=9583679;
RA Arama E., Yanai A., Kilfin G., Motro B.;
RT "Murine NIMA-related kinases are expressed in patterns suggesting
RT distinct functions in gametogenesis and a role in the nervous
RT system.";
RL Oncogene 16:1813-1823(1998).
RN [3]
RP SEQUENCE FROM N.A., AND TISSUE SPECIFICITY.
RC TISSUE=Brain;
RX MEDLINE=98096235; PubMed=9434622;
RA Tanaka K., Parvinen M., Nigg E.A.;
RT "The in vivo expression pattern of mouse Nek2, a NIMA-related kinase,
RT indicates a role in both mitosis and meiosis.";
RL Exp. Cell Res. 237:264-274(1997)
CC -!- FUNCTION: PROTEIN KINASE THAT IS INVOLVED IN MITOTIC REGULATION.
CC MAY HAVE A ROLE AT THE G2-M TRANSITION. MAY ALSO PLAY A ROLE IN
CC MEIOSIS.
CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
CC -!- SUBUNIT: Interacts with TERP1 (By similarity).
CC -!- SUBCELLULAR LOCATION: Nuclear.
CC -!- TISSUE SPECIFICITY: MOST ABUNDANTLY EXPRESSED IN TESTIS. LOW
CC LEVELS FOUND IN MID-GESTATION EMBRYO, OVARY, PLACENTA, INTESTINE,
CC THYMUS AND SKIN. WITHIN THE TESTIS, EXPRESSION RESTRICTED TO GERM
CC CELLS WITH HIGHEST LEVELS DETECTED IN SPERMATOCYTES AT PACHYENE
CC AND DIPLOTENE STAGES. ALSO EXPRESSED IN MEIOTIC PACHYTENE OOCYTES.
CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA
CC subfamily.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC EMBL; U95610; AAB67973.1; -
```

```
DR EMEL; AF013166; AAC35393.1; -
DR ENBL; AF007247; AAB70470.1; -
DR HSP; P24941; ICRP.
DR MGD; MGI:109359; Nek2.
DR InterPro; IPR000719; Prot_kinase.
DR InterPro; IPR008271; Ser_thr_pkin_AS.
DR InterPro; IPR002290; Ser_thr_pkinase.
DR Pfam; PF00069; pkinase; 1.
DR ProDom; PD000001; Prot_kinase; 1.
DR SMART; SM00220; S_TKC; 1.
DR PROSITE; PS00107; PROTEIN KINASE ATP; FALSE_NEG.
DR PROSITE; PS00108; PROTEIN KINASE ST; 1.
DR PROSITE; PS00111; PROTEIN KINASE_DCM; 1.
KW Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
KW Nuclear protein; Phosphorylation; Cell cycle; Cell division; Meiosis.
FT DOMAIN 8 271 PROTEIN KINASE.
FT NP_BIND 14 22 ATP (BY SIMILARITY).
FT BINDING 37 37 ATP (BY SIMILARITY).
FT ACT_SITE 141 141 BY SIMILARITY.
FT MOD_RES 175 175 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
SQ SEQUENCE 443 AA; 51307 MW; DE09565C378307E1 CRC64;

Query Match 14.7%; Score 532.5; DB 1; Length 443;
Best Local Similarity 31.9%; Pred. No. 7.5e-22;
Matches 146; Conservative 92; Mismatches 171; Indels 49; Gaps 15;

Qy 1 MDKYDVKAIGQAFKAYLAKGSKDKHCVIKEIFEKMPIQEKEASKEVILLERMKH 60
Db 5 VEDYEVLSHIGTGYRCQKIRKSDGKILVWKELDYGMTVEVKQMLVSEVNLRELKH 64
Qy 61 PHIVAFNSF--QENGRLPFVWEYCDGGLMKRN---RQCVLSEQILGWFOISLG 115
Db 65 PHIVRYDRIIDRTNTTLYVWEYCEGGLASVISKGTDRQYL--EESFVLRVMTQLTIA 123
Qy 116 LKHIDRK-----ILHRDIKAQNIIFLSKNGVMYAKLGDGFIARVLNNSMELARTCIGTPY 170
Db 124 LKECHRSDGHTVLRDLKPNVFLDSKHNV-KLGDGFLARILNHDTSFAKTFVGTYY 182
Qy 171 LSPETCONKPYNNKTDIWSLCVAYELCTLKHPPFEGNNLOQLVLKICQAHFAPISPGFSR 230
Db 183 MSPEQMSCLSYNEKSDIWSLACLLYELCALMPFTAFNQKELAGIRGFRFIRPYRSD 242
Qy 231 ELHSLISQLFQVSPDRPSINSILKRPFLNLPKYLTPVTOEBSFHMLI---CRAGAP 287
Db 243 GLNDLITRMFLDKYHRPSVEILSP-----LIADVAEQRNLBERGRSRSEP 293
Qy 288 AS-RHAGKYVQKCIQKRVFOGKCPP-RSRISVPKRNAILHRNEWPPAGAKARSII-- 343
Db 294 SKLPDSSPVLSELKESQLQDRQALRAREDIQKERELCIRERLAEDKLARAESLMK 353
Qy 344 --XMERPKIAVCGHYDYVYACLDM---LERRAKPSYHPIPOENT---GYEDYQGE 393
Db 354 NVSLKHEHLLCLAGG-----PELDLPSSAMKKKVH---FHGESKENTARSENSESYLAK 405
Qy 394 TRHGSPSPQWPAEYLQKFEA---QQYKLVKEQKGLR 428
Db 406 SKCRDLKXELHAAQLRAQALADIENKYQLKSRQILGMR 443

RESULT 13
NEK9_MOUSE
ID NEK9_MOUSE STANDARD; PRT; 984 AA.
AC Q8KLR7; O833P1;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Serine/threonine-protein kinase Nek9 (EC 2.7.1.37) (Nima-related
DE protein kinase 9).
GN NEK9 OR NERCC.
OS Mus musculus (Mouse).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
OX NCBI_TaxID=10090;
```

RN SEQUENCE FROM N.A.
 RC STRAIN=BALB/c; TISSUE=Heart;
 RA Joch M.R., Mandalino C., Scime A., Whyte P.;
 RT "A novel Nima related protein kinase containing an RCC1-like domain."
 RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
 RN
 RP SEQUENCE OF 759-984 FROM N.A.
 RC TISSUE=Breast tumor;
 RX MEDLINE=22388257; PubMed=12477932;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Sherman C.M., Schuler G.D.,
 RA Altschul S.F., Zerborg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
 RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M.J., Udell T.B., Toshikiyuki S., Carninci P., Schetz T.E.,
 RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalón D.K., Muny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahey J., Helton E., Kettman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywinski M.I., Skalska U., Smalhus D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Marra W.A.;
 RA "Generation and initial analysis of more than 15,000 full-length
 RT human and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 CC -!- FUNCTION: Pleiotropic regulator of mitotic progression,
 CC participating in the control of spindle dynamics and chromosome
 CC separation. Phosphorylates different histones, myelin basic
 CC protein, beta-casein, and BICD2. Phosphorylates histone H3 on
 CC serine and threonine residues and beta-casein on serine residues
 CC (By similarity).
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- SUBUNIT: Homodimer. Binds to Ran GTPase. Has a greater affinity
 CC for Ran-GDP over Ran-GTP. Interacts with NEK6, NEK7 and BICD2 (By
 CC similarity).
 CC -!- SUBCELLULAR LOCATION: Cytoplasmic (By similarity).
 CC -!- DOMAIN: Dimerizes through its coiled-coil domain (By similarity).
 CC -!- PTM: Autophosphorylated on serine and threonine residues (By
 CC similarity).
 CC -!- SIMILARITY: Belongs to the Ser/Thr family of protein kinases. NIMA
 CC subfamily.
 CC -!- SIMILARITY: Contains 6 RCC1 repeats.
 CC
 CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC
 DR EMBL; AJ489828; CAD34025.1; .
 DR EMBL; BC024926; AAH24926.1; ALT_INIT.
 DR MGD; MGI:2387995; Nek9.
 DR InterPro; IPR000719; Prot_kinase.
 DR InterPro; IPR000408; Reg_Chrcondens.
 DR InterPro; IPR008271; Ser_thr_pkin_AS.
 DR InterPro; IPR002290; Ser_thr_pkinase.
 DR InterPro; IPR001245; Tyr_pkinase.
 DR Pfam; PF00069; pkinase; 1.
 DR Pfam; PF00415; RCC1; 2.
 DR PRINTS; PR00109; TYRKINASE.
 DR ProDom; PD000001; Prot_kinase; 1.
 DR SMART; SM00220; S_TKC; 1.
 DR SMART; SM00219; TyrK; 1.
 DR PROSITE; PS00107; PROTEIN_KINASE_ATP; FALSE_NEG.
 DR PROSITE; PS00111; PROTEIN_KINASE_DOM; 1.
 DR PROSITE; PS00108; PROTEIN_KINASE_ST; 1.

DR PROSITE; PS00625; RCC1_1; FALSE_NEG.
 DR PROSITE; PS00626; RCC1_2; FALSE_NEG.
 DR PROSITE; PS0012; RCC1_3; 6.
 KW Transferase; Serine/threonine-protein kinase; ATP-binding; Mitosis;
 KW Phosphorylation; Cell cycle; Cell division; Repeat; Coiled coil.
 FT DOMAIN 52 308 PROTEIN KINASE.
 FT NP_BIND 58 66 ATP (BY SIMILARITY).
 FT REPEAT 388 444 RCC1 1.
 FT REPEAT 445 498 RCC1 2.
 FT REPEAT 499 550 RCC1 3.
 FT REPEAT 551 615 RCC1 4.
 FT REPEAT 616 668 RCC1 5.
 FT REPEAT 669 726 RCC1 6.
 FT DOMAIN 732 896 INTERACTION WITH NEK6 (BY SIMILARITY).
 FT DOMAIN 772 893 SER/THR/PRO-RICH.
 FT DOMAIN 896 945 COILED COIL (POTENTIAL).
 FT DOMAIN 752 767 POLY-GLY.
 FT ACT_SITE 176 176 BY SIMILARITY.
 FT BINDING 81 81 ATP (BY SIMILARITY).
 FT MOD_RES 210 210 PHOSPHORYLATION (AUTO-) (BY SIMILARITY).
 FT CONFLICT 929 929 S -> F (IN REF. 2).
 SQ SEQUENCE 984 AA; 107083 MW; F64A26AA0466749D CRC64;
 Query Match 14.7%; Score 532.5; DB 1; Length 984;
 Best Local Similarity 36.8%; Pred. No. 1.8e-21;
 Matches 118; Conservative 59; Mismatches 121; Indels 23; Gaps 5;
 QY 4 YDVKAIQGAGFGKAYLAKGSKSCVCVKEINFERMKIQKEASKEVILLKMKHPNI 63
 DB 52 YIPRVLRGAFGEATLYRRTEDDSLVWKEVDLTLSKERDALNEIVILALQHDNI 111
 QY 64 VAFNSFQENGLFIWVEYCDGDLMKRINRQGVLFSEDCILGFWFOISLGLKHIDRK 123
 DB 112 IAYNHFMNDNTLLIELEYCNGNLYDKILRQDKJFEEMVWVYLFQIVSAVSCIHKAG 171
 QY 124 ILHRDIKAQNIFLSKNGWAKLDFGIARVNNNSMELARTCIGTPYVLSPEICQNKPYNN 183
 DB 172 ILHRDIKTLNIFLTKANLI-KLGDYGLAKLNSEYSAETLVGTPIYMSPELQGVKYNF 230
 QY 184 KTDIWSLGVLYELCTLKHFPENNLQOLVLIKCOAHFA--PISPGFSRLHSLISQLFQ 241
 DB 231 KSDIWAYGCVIFELLTLKRTFDATNPLNLCVKIVQIGIRAMEVDSSQVSLQLVHACL 290
 QY 242 VSPRDRPSINSILKRPFLNLI PKYLTPEVIQEFSHMLICRAGAPASRAGHVQCKI 301
 DB 291 QDEQRPAADALDLFLLR-----TRRMEKEVTLL-----NAPTKRPSRSTVEAPI 339
 QY 302 QKVRFQ-----GKCPPR 313
 DB 340 AVTSTRSEVYVWGGGKSTPQ 360
 RESULT 14
 ID NEK9_XENLA STANDARD; RET; 944 AA.
 AC Q7ZCC8;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 10-OCT-2003 (Rel. 42, Last annotation update)
 DE Serine/threonine-protein kinase Nek9 (EC 2.7.1.37) (Nima-related
 DE protein kinase 9) (XNEK9) (Nercc1 kinase).
 GN NEK9.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae; Pipidae;
 OC Xenopodinae; Xenopus.
 OX NCBI_TaxID=8355;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Roig J., Avruch J.;
 RL Submitted (APR-2003) to the EMBL/GenBank/DBJ databases.
 CC -!- CATALYTIC ACTIVITY: ATP + a protein = ADP + a phosphoprotein.
 CC -!- SUBUNIT: Homodimer (By similarity).

Db 86 HEENDRLIIVMEFADSGNLDEQI-KLRGSGDARYQOEHEALFLQLCLALDYIHSHKML 144
QY 126 HRDIKAQNIFLSKNGWAKLGDGFIARVLNNSME--LARTCIGTPYXLSPEICONKPYNN 183
Db 145 HRDIKSANVLITSTGLV-KLGDGFGSHQYEDTVSGVASTFCGTPYLLAPELWNNKRYNK 203
QY 184 KTDIWSLGCVLVELCTLKHPPGNNLQOLVLKICOAHFAPISPGFSRELHSLISQLFQVS 243
Db 204 KADVWSLGLVLYEIMGKMKPFASNLKGLMSKVLAGTYAPLPDGSFSSEFKRWVDGILVAD 263
QY 244 PDRPSINSILKRPELENIIPXYL-----TPEVIOEEFHHMLICRAGAPASR- 290
Db 264 PNDPSVREIFQIPYINKGLFVQALKKNERISDSVKEVLVTQVSEILSSEVSFPAHRF 323
QY 291 -----HAGKVYQ----KCKIQKVRF-----QGK 309
Db 324 LVSQINYDVTHRGHYNKLGNGGKSKMKRFLQIVRGQLILTDDEGNNPKGINLEQVQGA 383
QY 310 CP-PRS 314
Db 384 CVPHS 389

Search completed: October 8, 2004, 17:08.08
Job time : 19 secs

GenCore version 5.1.1.6
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: October 8, 2004, 17:05:06 ; Search time 21 Seconds
(without alignments)
3128.515 Million cell updates/sec

Title: US-09-940-921B-2
Perfect score: 3615
Sequence: 1 MDKYDVIRKAGQAFKAYL.....SVSRQEGTKDPYSPVLILM 683
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 78: *
1: Pir1: *
2: Pir2: *
3: Pir3: *
4: Pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1106.5	30.6	774	2 J25284	protein kinase nek
2	665.5	18.4	792	2 J2122	protein kinase (EC
3	660	18.3	841	1 I78885	serine/threonine-s
4	589	15.7	357	2 T29771	hypothetical prote
5	539.5	14.9	445	2 G01452	NIMA-like protein
6	521.5	14.4	431	2 T11854	protein kinase (EC
7	517.5	14.3	699	2 A43734	probable protein k
8	511	14.1	722	2 T37970	probable G2-specif
9	471.5	13.0	338	2 J27838	Nek6 protein kinas
10	464.5	12.8	779	2 A51777	NIMA-like protein
11	457.5	12.7	941	2 T49136	protein kinase-lik
12	424	11.7	294	2 T21075	hypothetical prote
13	423	11.7	435	2 J23580	probable protein k
14	415.5	11.5	200	2 B95887	hypothetical prote
15	410	11.3	648	2 T47988	serine/threonine-p
16	403.5	11.2	925	2 A55748	protein kinase (EC
17	402.5	11.1	1895	2 T15881	hypothetical prote
18	397	11.0	576	2 S22127	protein kinase pol
19	396.5	11.0	114	2 T38224	protein kinase/thr
20	393.5	10.9	465	2 B55748	protein kinase (EC
21	392	10.8	1142	2 S59359	GIN4 protein - yea
22	391	10.8	658	2 T39500	serine/threonine-s
23	388	10.7	705	2 A48144	protein kinase CDC
24	387.5	10.7	1051	1 JW0051	serine/threonine-s
25	382.5	10.6	982	2 T18576	serine-threonine k
26	382	10.6	461	2 T14822	probable serine/th
27	381	10.5	257	2 T50802	probable protein k
28	379	10.5	445	2 G54797	serine/threonine p
29	379	10.5	682	2 A44493	serum-inducible ki

30 379 10.5 752 1 A32571 ribosomal protein
31 379 10.5 856 2 T43631 serine/threonine k
32 378 10.5 490 2 S47946 protein kinase hom
33 376 10.4 733 2 A57459 ribosomal protein
34 374.5 10.4 897 2 S61137 probable membrane
35 373.5 10.3 608 2 G96575 probable MEK kinas
36 372 10.3 735 2 A53300 ribosomal protein
37 371.5 10.3 363 2 T08542 mitogen-activated
38 371.5 10.3 363 2 T51735 protein kinase (EC
39 371 10.3 819 2 A53714 protein kinase (EC
40 370.5 10.2 740 2 T24340 hypotheical prote
41 370.5 10.2 797 2 T23927 hypotheical prote
42 369.5 10.2 1001 2 T17365 serine/threonine p
43 368.5 10.2 735 2 T51901 ribosomal protein
44 368.5 10.2 836 2 B96716 probable serine/th
45 368.5 10.2 1558 2 T29253 hypotheical prote

ALIGNMENTS

RESULT 1

S25284
C:Species: Mus musculus (house mouse)
C>Date: 17-Apr-1993 #sequence_revision 17-Apr-1993 #text_change 24-Sep-1999
C/Accession: S25284
R/Lewis, K.; Mizzen, L.; Motro, B.; Ben-David, Y.; Bernstein, A.; Pawson, T.
EMBO J. 11, 3521-3531, 1992
A>Title: A mammalian dual specificity protein kinase, Nek1, is related to the NIMA cell
A/Reference number: S25284; MUID:93010942; PMID:1382974
A/Accession: S25284
A>Status: not compared with conceptual translation
A/Molecule type: mRNA
A/Residues: 1-774 <DET>
A/Cross-references: GB:S45828; NID:g256854; PIDN:AB23529.1; PID:g256855
C/Genetics:
A/Gene: nek1
A/Superfamily: unassigned ser/Thr or Tyr-specific protein kinases; protein kinase homolc
C/Keywords: ATP; phosphotransferase
F:2-258/Domain: protein kinase homology <KIN>
F:10-18/Region: protein kinase ATP-binding motif

Query Match 30.6%; Score 1106.5; DB 2; Length 774;
Best Local Similarity 35.2%; Pred. No. 8.1e-37;
Matches 276; Conservative 121; Mismatches 207; Indels 181; Gaps 23;
QY 1 MDKYDVIRKAGQAFKAYLAKGKSDSKHCIVIKENIFKPIQKKEAKKEVILLERMKH 60
Db 1 MEKYVRLQKIGESGFGKAVLVKSTEDGRHYVIKEINISMSDKERQESREAVLANMKH 60
QY 61 PNIVAFNSFQENGRLPIVMVEYCDGDLMKRINRQGVLFSEDIQILGWFYQISLGLKHIH 120
Db 61 PNIVQYKESFEENGSLYIVMDYCEGGDLFRINAQKALFQEQILDFVQICLALKVHV 120
QY 121 DRKILHRDIKAQNIFFLSKNGWAKLGFQGIARVLNNSMELARTICGTPPYLSPEICQKPK 180
Db 121 DRKILHRDIKSNIFLTKDGV-QLGDFGIARVLNVELARTICGTPPYLSPEICENKP 179
QY 181 YNNKTDIWSLGVLYELCTLKHPEGNLQQLVLKICQAHFAPISCFPSRELHSLISQLF 240
Db 180 YNNKSDIWAIGCVLYELCTLKHFAEAGNMKNVLKISGSPFPVSPHYSYDLKSLLSQLF 239
QY 241 QVSRDRPFSINSILKRPFLENIPKYLTPTEVIEEFSSHMLICRAG----- 285
Db 240 KRNPRDRPSVNSILEKGFIAKRIEKFSLPQIAEEFCLKLSKFGPQPLGPKRPASQGV 299
QY 286 -----APASRHAGKVQCKQKVFQKCPSPRIS--VPKRNAILHRNEWRP 333
Db 300 SSFVPAQKITKPAKYGVPITYKKGKLEKPPKPKHQAHQIPVKK---MNSGEERK 356
QY 334 PAG-AQKASIKMIER----- 348

```

Db 357 KMSEBAKKRRLEFTEKEKKQDQIRFLKAEQMKRQKQRLERINRAREQGRNVLRAAG 416
QY 349 -----PKIAAVCGHYDYVYQAQDMLRE-BAH-----KPSYHP--IPEO 383
Db 417 SGEVKAFFGIGAVSPSPCSRPGQYEHYHAIFDQMRADNEARWKGGIYGRWLPER 476
QY 384 NTG---VEDYGOETRHGSPSPQWPAEYLQRFKFEAQYKLYKE-----KQLGLRPS 431
Db 477 QXGHIAVERANQ-----VEEFLQRKEAMQKARAEAGHVYVLAURLQIRLQ--- 522
QY 432 AEPNYNQREL-----SNGEPRPQELPFRKNEKMQEYKQLEERIQYHN 479
Db 523 ---NFERQIQIKAKURGENKEADGFKQGEAT-EETDMR---LKKWESLKAQTNARAVALK 575
QY 480 DMKEIRKMGKRPENSKISHKTYLVKSNLVPVHODASE-GEAPVQDIEKDKMLQNT 538
Db 576 EQLERKREAYEREKVVWEHLVARVKSVDVPLPLELLETGSP----- 619
QY 539 KESKNPEQYKAKGVKFEINLDCISDENILQEEAMDIENETLT-----FEDGMK 590
Db 620 --SKQVFPVSVTSALKKEVGLDGLTD-----TOBEEMKENSATSKREILRLNENLK 673
QY 591 FREYECVKEHGDTYTKAFELKHCPEAGFSTOTVAAVGNRRQWDGAGAPOTLQWMAVADIT 650
Db 674 AQEDEKEKQH--HSGSCETVGHKDBREYETEN-AISSDKKEMWGG-QLVPLDAVLTDT 729
QY 651 STCPT 655
Db 730 SPSAT 734

RESULT 2
JC7122
Protein kinase (EC 2.7.1.37) 2 - mouse (strain balb/c)
N/Alternate names: serine (threonine) protein kinase
C:Species: Mus musculus (house mouse)
C:Date: 23-Mar-2001 #sequence_revision 23-Mar-2001 #text_change 31-Mar-2001
C:Accession: JC7122
R:Hayashi, K.; Igarashi, H.; Ogawa, M.; Sakaguchi, N.
Biochem. Biophys. Res. Commun. 264, 449-456, 1999
A:Title: Activity and substrate specificity of the murine STK2 serine/threonine kinase
A:Reference number: JC7122; MUID:20001940; PMID:10529384
A:Accession: JC7122
A:Status: preliminary
A:Molecule type: mRNA
A:Residues: 1-792 <HAY>
A:Cross-references: GB:AJ223071; NID:94138208; PID:94138209
C:Genetics:
C:Superfamily: human serine/threonine-specific protein kinase STK2; protein kinase homol
C:Keywords: phosphotransferase

Query Match 18.4%; Score 655.5; DB 2; Length 792;
Best Local Similarity 25.3%; Pred. No. 1.6e-19;
Matches 189; Conservative 137; Mismatches 213; Indels 209; Gaps 23;
QY 4 YDVIRKAGOCAGFKAYLAKGSDSKHCVIKINFEKMPIQEKASKKEVILLERKHPNI 63
Db 6 YCYMRVVGSGSYGEVTLVKHREDGQYVYKLNLRNASSRERRAAEQEQLLSQLKHPNI 65
QY 64 VAFNFSQ-ENGRLEFVMEYCDGGLMKRINRQGVLSFSDQILGNFVQISLGLKHDR 122
Db 66 VTYKESWEGDGLLYVNGFCGGDLRYKLKBQGLLPESQVVEFWQIAVALQYLHEK 125
QY 123 KILHRDIKAQNTFLKNGWAKLGFGIARVLNNSMELARTCIGTPYVLSPEICQNKPN 182
Db 126 HLHRDLTKQVLFRTNII-KVGDGLIARVLNHNHGDWASTLIGTPYVMSPELSNKPYN 184
QY 183 NKTDTWSLGCVLVELCTLKHPPGNNLQVLVKICQAHFAISPGFSELHSLISQLFQV 242
Db 185 YKSDVWALGCCVYEMATLKHAFNAKDMNSLVRIIEGKLPPMPKVYSTELAEIIRTLSR 244
QY 243 SPRDRPSINSILKRPFLNLI PKYLTPTVQIEPFSHMLIC----- 282

```

```

Db 245 RPEERDSVRISILQPIKHHISLFLEATKAKTSKNNVKNCDSRAPKAAVAVSRKESNTD 304
QY 283 -----RAGAPASHAGKVVQVKCKIKVRFG 308
Db 305 VIHYQPSRSEGAHVMGEDKCLSQEKFPVDIGPLSPASLEGHTCK-----QDMNNTG 357
QY 309 K-CPPRSRISVPIKRNAILERNEMPPAGAAQAKARISKIE--RPKIAAVCGHYDYVYQAQ 365
Db 358 ESCATISIRINIDIL-----PAERRDSANAGVVGESQPCHVDAADEVD---SQC 402
QY 366 DMLRRRAH-----KPSYHP--IPEOENT-GVEDYGOETRHGSPSPQWPAEYLQRFKFEAQ 416
Db 403 STSQEKERLQGNTKSDQGNLLPRSSDGGGEGSELVKPLYPN-----KDKQPDQDQ 457
QY 417 YKLKVKQLGLRP-----SSAEPNPN-----ORQLSRNGEERPFQELPPR----- 457
Db 458 VTGIIENQDSIHPRSQPHSSMSEPSLSRORRQKREQTAHSGTKSQFQELPPLPLSPYPG 517
QY 458 -----KNE-----MKQEYWKQLEE----- 472
Db 518 IGVKDIATQQDNGNQGGPVAGCVNSRTSTASAKDRPLSARERRRLKQSOEMLPSGP 577
QY 473 -----IRQVHNDMKIIRKMGKREPENSKISHK 501
Db 578 AVQRTPSAVEPLKQBEDQPIPAQRPSSDCSITQXNHTLPREKELMHGLSED-ELSSS 636
QY 502 TYLVKSNLVPVHODASEGAPVQDIEKDKMLQRLQ-----NTKES-----KNPEQYKAK 551
Db 637 T8STDKSD---GDSREGKSHTNEM-KDLVQLMTQTLREAKESCEDLQVLNPGSEFRLH 691
QY 552 KGVKFEINLDCISDENILQEEAMDIP 579
Db 692 RKYRDTLVHLGKVAEB---VEPHCTELP 716

RESULT 3
I78985
serine/threonine-specific protein kinase (EC 2.7.1.-) STK2 - human
C:Species: Homo sapiens (man)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 23-Sep-1999
C:Accession: I78985
R:Levedakou, E.N.; He, M.; Baptist, E.W.; Craven, R.J.; Cance, W.G.; Welcsh, P.L.; Simmor
Oncogene 9, 1977-1988, 1994
A:Title: Two novel human serine/threonine kinases with homologies to the cell cycle regu
A:Reference number: I58396; MUID:94268938; PMID:8208544
A:Accession: I78985
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-841 <RES>
A:Cross-references: GB:I20321; NID:G348244; PIDN:AAA36658.1; PID:G348245
C:Genetics:
A:Gene: GDB:STK2
A:Cross-references: GDB:374125
A:Map position: 3p21.1-3p21.1
C:Superfamily: human serine/threonine-specific protein kinase STK2; protein kinase homol
C:Keywords: phosphotransferase
F:4-261/Domain: protein kinase homology <KIN>

Query Match 18.3%; Score 660; DB 1; Length 841;
Best Local Similarity 25.3%; Pred. No. 2.8e-19;
Matches 200; Conservative 141; Mismatches 238; Indels 212; Gaps 29;
QY 4 YDVIRKAGOCAGFKAYLAKGSDSKHCVIKINFEKMPIQEKASKKEVILLERKHPNI 63
Db 6 YCYLRVVGSGSYGEVTLVKHREDGQYVYKLNLRNASSRERRAAEQEQLLSQLKHPNI 65
QY 64 VAFNFSQ-ENGRLEFVMEYCDGGLMKRINRQGVLSFSDQILGNFVQISLGLKHDR 122
Db 66 VTYKESWEGDGLLYVNGFCGGDLRYKLKBQGLLPENQVVEFWQIAVALQYLHEK 125
QY 123 KILHRDIKAQNTFLKNGWAKLGFGIARVLNNSMELARTCIGTPYVLSPEICQNKPN 182

```

Db 126 HILHDLKTONVFLRTNII-KVGDGLGARTVLNENHCDMASTLIGTPYVMSPELFSNKPYN 184
QY 183 NKTIDWSLGVLYELCTLKHPPGEGNNLQOLVLIKQAHFAPISQFSLHSLLSQLPQV 242
Db 185 YKSDVWALGCCVYEWATLKHAFNAKMDMSLVYRIEGLPAMPDRDYSPELAELRTMLSK 244
QY 243 SPRDPSINSILKRPFLNLPKYLTPVQI-----EEFHHMLICRAGAPAS 289
Db 245 REEPSPVRSILRQYIKRQISFFLEATKIKTSKNINKNQDSQSKPPATVV---SGEAS 301
QY 290 RHAGKVQCKIQK-----VFQKC--PPRSISVPIKRNAILHRNEWPPAGAOKAR 341
Db 302 NH--EVIHPQLPSLSSGSGTYINGEGKCLSQKPRASGLKSPASL-----KAH 347
QY 342 SIK--MIERPKAAVCGHYDYVYAOLDMLRRA-----HKPSYHPIPOENTQVE 388
Db 348 TKQDLSNTELATISS-----VNIDILPAKGRDSVSDGFQENQPRYLNASNELGGIC 401
QY 389 DYGO-----ETHRGSPS---QW-----PAEYLQKRFQAQYKXK----- 420
Db 402 SISQVEEEMLDQNTKSSAQENLIPWSSDIVTGEKNEPVKPLQPLIKEQKPKQSLALS 461
QY 421 -----VEKQLGRP-- 429
Db 462 PKLECSGITLAHNSRLILGSSDPSASASRVACITGVCHHAQDQVAGECIIIEKQRIHPDL 521
QY 430 -----SSAEPNPN-ORQELRSNGEPR-----FQELPFR-----KN 459
Db 522 QPHNSGSPSLSRQKREKRETEHRGEKQVRDLFAFQSPPRFLPSHPIVGVKDVTS 581
QY 460 ENKEQYWKQ-----LEIRQYHNDKX-----IRKNGREBENSISKHYLYVKKS 508
Db 582 TOKAENQRRVVTGVSRSRSESSKDRPLASERERRLKQSQEEMSSGSPS---VKA 638
QY 509 NLPVHQDASEGAPQD-----TEKDLKQML-----QNTKESKNPEOKY 548
Db 639 SILV---ACPGKQEDQPLPARRLSSDCSVTQERKQIHCLSEDLSSSTSDTKSDGDY 695
QY 549 KAKGVKFEINDKICSDENIL--QEEBAMDIP--NETLTFEDGKMFKEVBCVKEHGDYT 604
Db 696 GEGKQGTNEINALVQMTQTLDKSDCEDVPVNPVSEPKLHRKYRDTLIL--HGRVA 753
QY 605 DKAFELHCPPE 615
Db 754 EEA-EELHFKPE 763

RESULT 4
T29771
hypothetical protein ZC581.1 - Caenorhabditis elegans
C:Species: Caenorhabditis elegans
C>Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
C:Accession: T29771
R:Waterston, B.; Gattung, S.; Le, T.T.
A:Description: The sequence of C. elegans cosmid ZC581.
A:Reference number: Z20682
A:Accession: T29771
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-357 <WAT>
A:Cross-references: EMBL:AF003134; PIDN:AA54139.1; GSPDB:GN00019; CBSP:ZC581.1
A:Experimental source: strain Bristol N2; clone ZC581
C:Genetics:
A:Gene: CBSP:ZC581.1
A:Map position: 1
A:Introns: 31/3; 81/1; 120/3; 186/1; 226/3; 260/1; 288/3
C:Superfamily: Ca2+/calmodulin-dependent protein kinase I; protein kinase homology
Query Match 15.7%; Score 569; DB 2; Length 357;
Best Local Similarity 40.9%; Pred. No. 4.9e-16;
Matches 113; Conservative 59; Mismatches 90; Indels 14; Gaps 4;

QY 1 MDKYDVIKAGCGAFKAYLAKGKSDSKH--CVIKENFEKMPIQEKASKKEVILLEKM 58
Db 1 MDYKRVVVGAGFAGVGCWLCRGKNDASHQKVIILKINTHGTKEENSIQSEVNLKKV 60
QY 59 KHNIVAFNPSFQENGLFTVMEYCDGGDLMKRINQGV-----LFSDDLGLGFVQ 111
Db 61 QHPLITGYIDSFTMDNQLGIVMCYAGSGTLERLINDQRAIKDSNNMREYFPEKTVLDYFTQ 120
QY 112 ISLGLKHIDRKILHRDIKAQNIFLSKNGMAVAKLGDGFIARVLNNSMELARTCIGTPYYL 171
Db 121 ILTALNHHQKXIVHRDLKPNILMNRKTVLKLSDFGISKEL-GTKSAASTVIGTPNYL 179
QY 172 SPEICQNKPNYKNTDWSLGCVLVELCTLKHPPGEGNNLQOLVLIKQAHFAPISQFSSRE 231
Db 180 SPEICSRPNYQKSDWMSLGCVLVELCTLKHPPGEGNNLQOLVLIKQAHFAPISQFSSRE 239
QY 232 LHSLSIQLFQVSPRDRPSINSILKRPFLNLPKYL 267
Db 240 VKMLVENLLKTHTKRPDVSQLSDP-----LVLPYL 271

RESULT 5
G01452
NIMA-like protein kinase 1 - human
C:Species: Homo sapiens (man)
C>Date: 21-Dec-1996 #sequence_revision 06-Jun-1997 #text_change 24-Sep-1999
C:Accession: G01452
R:Lu, K.P.
A:Submitted to the EMBL Data Library, June 1994
A:Reference number: G07172
A:Accession: G01452
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-445 <JUX>
A:Cross-references: EMBL:U11050; NID:G507874; PIDN:AAA19558.1; PID:G507875
C:Genetics:
A:Gene: NLK1
C:Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homology
C:Keywords: ATP
F:6-271/Domain: protein kinase homology <KIN>
F:14-22/Region: protein kinase ATP-binding motif

Query Match 14.9%; Score 539.5; DB 2; Length 445;
Best Local Similarity 29.9%; Pred. No. 8.6e-15;
Matches 141; Conservative 88; Mismatches 167; Indels 75; Gaps 13;

QY 2 DKYDVIKAGCGAFKAYLAKGKSDSKHCHVCVIKENFEKMPIQEKASKKEVILLEKM 61
Db 6 EDYEVLYTIGTGYGRCQKIRKSDGKILVWKELDYGSMTAEAKQMLVSEVNLRELKHP 65
QY 62 NIVAFNSF--QENGLFIWMEYCDGGDLMKRI---NRQGVLFSEDDILGWTFVQISGL 116
Db 66 NIVRYDRIIDRTNTLYIWEYCEGGDLASVITKGTKEQYL-DEEFVLVMTQLTLAL 124
QY 117 KHIHDRK-----ILHRDIKAQNIFLSKNGMAVAKLGDGFIARVLNNSMELARTCIGTPYYL 171
Db 125 KECHRRSDGGHTVLRDLKPNVFLDQKQNV-XLGGDFGLARILNHDTSFAKTFVGTPIYM 183
QY 172 SPEICQNKPNYKNTDWSLGCVLVELCTLKHPPGEGNNLQOLVLIKQAHFAPISQFSSRE 231
Db 184 SPEQNMNMSYNEKSDIWSLGCVLVELCALMPPTATFSQKELAGIKREGKFRPIRYVSD 243
QY 232 LHSLSIQLFQVSPRDRPSINSILKRPFLNLPKYL 267
Db 244 LNEIITRMLNLKDYHRSVEILENPLIADL-----VADEQRNLSRRGRQLGPE 294
QY 288 ASRHAGKVQCKIQKQVRFQKGCPPRRSRISVPDKRNAILHRNEWPPAGAOK----- 339
Db 295 KSDSSPVLSELKKEIQK-----ERELAKAREERLEQKQCELCVRLA 341
QY 340 -----ARSKITERPKIAAVCGHYDYVYAQLDMLRRRAHPSYHPIPOENTQVED 389
Db 342 EDKLARAEENLLKNYSILKEREKFLSLASNPPELLNLPSSVKKVH---FSGESKENI---- 394

Biochem. Biophys. Res. Commun. 233, 753-758, 2002

A;Title: Identification and characterization of Nek6 protein kinase, a potential human A;Reference number: JC7838; PMID:22050088; PMID:12054534

A;Accession: JC7838

A:Molecule type: mRNA

A;Residues: 1-338 <HAS>

A;Cross-references: DBJ:AB026289

C;Comment: This protein, a novel member of NIMA (never in mitosis, gene A) related kinase playing a conserved central role in regulating mitotic chromosome condensation and asse C;Genetics:

A;Gene: nek6

A;Map position: 9q33-34

Query Match 13.0%; Score 471.5; DB 2; Length 338;
Best Local Similarity 38.5%; Pred. No. 3.1e-12;
Matches 99; Conservative 53; Mismatches 96; Indels 9; Gaps 6;

QY 4 YDVIKAGQAGFCAYLAKGSDSKHCVKEIN- FEKMPIQKEASKKEVILLENKMKHP 62
DB 70 FQIEKKIGRGQSEVYKATCLDRKTVALKKVQIFEMMDAKARQDCVKEIGLLKQLNHEN 129
QY 63 IVAFFNSFQENGLFTWMEYCDGGDLMKIN-- RQRGVLFSDQILGNWFQISLGLKHII 120
DB 130 IIKYLDSPFTEDNELNIVLEADAGLSQMIKVKFKQRLIPERTVMKYFVQJCSAVEHMH 189
QY 121 DRKILHRD;KAQNIPLSKNGWAKLGDGFIAGVNNMELARTCIGTPYVLSPICQNKP 180
DB 190 SRVWRD;IKPANVFITAGVV- KLGDGLGGRFFSETTAHSLVGTPTYMGSPERHENG 248
QY 181 YNNKTDIWSIGCVLYELCTLKHFFEGN-- NLQOLVLIKQAHFAPISPG--FSRELHSLII 236
DB 249 YNFKSDIWSIGCLLYEMAALQSPFYGDKNLFLSQCIRQCDDYPL- PGEHYSEKLRLELV 307
QY 237 SQLFQVSPRDRPSINSI 253
DB 308 SMCICDPHQRPDYGIV 324

RESULT 10

A57177

NIMA-like protein kinase - Neurospora crassa

C;Species: Neurospora crassa

C;Date: 10-Nov-1995 #sequence_revision 10-Nov-1995 #text_change 11-Aug-2003

C;Accession: A57177

R;Pu, R.T.; Xu, G.; Wu, L.; Vierula, J.; O'Donnell, K.; Ye, X.S.; Osmani, S.A. J. Biol. Chem. 270, 18110-18116, 1995

A;Title: Isolation of a functional homolog of the cell cycle-specific NIMA protein kinase A;Reference number: A57177; PMID:95355415; PMID:7629122

A;Accession: A57177

A;Status: preliminary; not compared with conceptual translation

A:Molecule type: DNA

A;Residues: 1-779 <PUA>

A;Cross-references: GB:142573; NID:g1040682; PIDN:AAA80145.1; PID:g1040683

C;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo C;Keywords: ATP

F;5-290/Domain: protein kinase homology <KIN>

F;13-21/Region: protein kinase ATP-binding motif

Query Match 12.8%; Score 464.5; DB 2; Length 779;
Best Local Similarity 25.5%; Pred. No. 1.2e-11;
Matches 163; Conservative 97; Mismatches 213; Indels 167; Gaps 24;

QY 2 DKYDVIKATQGAFCAYLAKGSDSKHCVKEINFEKMPIQKEASKKEVILLENKMKHP 61
DB 5 DKYELLEKTIHGHSFGFIIRKVRKADGMILCRKEISYLKMKSQEREOLHAEFSILSTRHP 64
QY 62 NIVAFPN--SFQENGLFTWMEYCDGGDLMKRI-NRQGVLFSDQILGN--FVQISLGL 116
DB 65 NIVGYVHREHLKATQDLHLVMEYCGNGDLGRVIRNLIKNNQYAESFV-WSIFSQLVITAL 123
QY 117 KHHI-----DRKILHRD;KAQNIPLSKNGWAKLGDGFI 150
DB 124 YRCHGYDPPVEKGTVLGLGSTARPPSPGGMTILRDLKPNVFLGDNVS- KLGDGFL 182

QY 151 AEVLNNSMELARTCTIGTPYLSPEIQNPKYNTDINSGLCVLYELCTLKHPEGNLQ 210
 Db 183 SKVM-OSHDFASTYVGTTFYSPMSPECAAEKTYLSDIWSLGCIIYELCAREPPFNATHY 241
 QY 211 QLVLTICQAHFAPISPGFSRELHSISQFQVSPDRPSINSILKRPPLLENLIPKYLTP 270
 Db 242 QLVQIKKEGKIAPLPSVSGELFATIKCLRVNPDRTATLLNP-----IVLMEKE 296
 QY 271 VIOEFSEMLICRAGAPASRHAGKVQCKKQKVFQKCPKRSRISVPIKRNAILHRNE 330
 Db 297 KEWFEFRTL-----RTKEETLNK-RIRELDSKLASLETESKSRASIDASL-RRE 345
 QY 331 WRPPAGAOKARSIMIERPKTAAVGHYDYVYQAQDMLRRRAHKPSYHPIQENTGVEDY 390
 Db 346 WE-----VKAR-----LEIDL-----VAQE----- 361
 QY 391 GOETHRGSPSPWPAEYLQKPEAQYKLVKEQL-----GLRP-----SSA 432
 Db 362 -----IESLQKPE-QEVOARVEAELOHGRGPFMFNSHGQCGFSSTAATL 406
 QY 433 BBNVQORQLRNGEPRFQELPFRKNEMKEQYWKQLEERQQVHNDMKIRKMGREP 492
 Db 407 VSDYNLSSVSGDGGPPSTTD-----ITDISAESTDGSITTKIPRTP 450
 QY 493 EENS KISHKTY--LVKSNLPVHDAS--EGEAPVODIEKDLKQMLQNTK--ESKNPEQKY 548
 Db 451 FHRA---QTYSSAPAESVLGTPMDIEMASPTIASLSLSPRMALTKAPTTPNPMIF 506
 QY 549 KAKKGKVEINLDKICSDENILQEEAMDIPNETLTPEDG 588
 Db 507 -----GPEPTSDKSNWEVPRETEMIDSG 530

RESULT 11
 T49136
 protein kinase-like protein - Arabidopsis thaliana
 N:Alternate names: protein F26G5.150
 C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Jun-2000 #sequence_revision 02-Jun-2000 #text_change 02-Jun-2000
 C:Accession: T49136
 R:P/Angelo, M.; Vezzi, A.; Modesto, D.; Pigazzi, M.; Valle, G.; Mewes, H.W.; Rudd, S.; I
 submitted to the Protein Sequence Database, April 2000
 A:Reference number: Z25017
 A:Accession: T49136
 A:Status: preliminary
 A:Molecule type: DNA
 A:Residues: 1-941 <DAN>
 A:Cross-references: EMBL:AL353814; GSPDB:GN00061; ATSP:F26G5.150
 A:Experimental source: cultivar Columbia; BAC clone F26G5
 C:Genetics:
 A:Gene: ARSP:F26G5.150
 A:Map position: 3
 A:Introns: 33/2; 55/3; 77/3; 93/2; 108/3; 135/3; 150/1; 166/3; 194/1; 210/3; 218/3; 237/3

Query Match 12.7%; Score 457.5; DB 2; Length 941;
 Best Local Similarity 23.4%; Pred. No. 2.7e-11;
 Matches 152; Conservative 121; Mismatches 235; Indels 141; Gaps 21;

QY 1 MDKYDVIRKAGGAFGKAYLAGKSDSKHCIVKEINFEKMPIOEKEASKKEVILLKMKH 60
 Db 5 MDQVELMEQIGRAGFAGAILVHHKAERKYYLVKTLARQTERCRRSARQHMSTIARVQH 64
 QY 61 PNIVAFNSFQNG--RLFTVMEYCDGDLMKRINRQKGLFSEQILGMFVOISLGLKHI 119
 Db 65 PIVFEKAWKEGVCIVITVCYEGGDVIELMKSNVYFPEKLCWFTQILLAVEYL 124
 QY 120 HPRKILHDKIAQNIFLSKNGMAKLGDFGIARVANNMELARTCTIGTPYLSPEIQN 179
 Db 125 HSNYVLRDLKSNIFLTAKQDV--RLGDFGLAKTL--KADDLTSSVVGTPNTMCPPELLADI 182
 QY 180 PYNNTDINSGLCVLYELCTLKHPEGNLQOLVLKIQAHFAPISPGFSRELHSLSL 239

Db 183 PYGFKSDIWSLGCIIYEMAAYRPAKAFDMAGLISK-----KSTH----- 222
 QY 240 FQVSPDRPSINS-----ILKRPFLLENLIPKY-----LTPE----- 270
 Db 223 -QGNVKEEPRVSAKRWASBILKHPYLOPYVQYRPTLSAASITPEKPLNSRGRSMAES 281
 QY 271 -----VIOEFSEMLICRAGAPASR-----HAGRVVQCKKIQKVR 305
 Db 282 QNSNSSEKDNFVSDKNIRYVVPNGNKVTETDSGFVDEEDILDHVQCSAENGNLQSV- 340
 QY 306 FQKCPKRSRISVPIKRNAILHRNWRP-----PAGAOKAASIMIERPKIAAVCVGHYDY 361
 Db 341 -SATKPDGHIKLPV-----HSQRPDVIQPRHPKTIIRNIMVLKEEKAENG----- 387
 QY 362 YAQDMLRRRAHKPSYHPIQENTGVEDYQETHRGSPSPWPAEYLQKPEAQYKLVKEQL-----AEYLQKFEAQ 416
 Db 388 -----SPMRNRRSPS--SVPTQKNV-----TPSKIPKLGDIASHSKTNASTPI 431
 QY 417 YKLKVEKQGLRPSAEFNNQORQLRNGEPRFQEL---PFRKNEMKEQYWKQLEBI 473
 Db 432 PPSKLASDSARTPGSPPKHMPVIDSSPKLRPNDRISPSPAKHEAEEA----- 482
 QY 474 RQYHNDMKEIRKMGREPEENSKISHKTYLVKK--SNLPVHDASEGEAPVQDIEKDLK 531
 Db 483 -----MSVRRQRTPTLPARTSLIAHQRQLGADISNMAKETAKLHPSVSESETNSH 537
 QY 532 QMELQNTKESKNPEQK---YKAKGKVEINLDKICSDENILQEEAMD 577
 Db 538 QSRVHASPVSTTPEPRKTSVGSAGKMQSESS--NGISSLSMQAFELCD 584

RESULT 12
 T21075
 hypothetical protein F19H6.1 - Caenorhabditis elegans
 C:Species: Caenorhabditis elegans
 C:Date: 15-Oct-1999 #sequence_revision 15-Oct-1999 #text_change 17-Mar-2000
 C:Accession: T21075; T21124
 R:McMurray, A.
 submitted to the EMBL Data Library, August 1995
 A:Reference number: Z19368
 A:Accession: T21075
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-294 <WIL>
 A:Cross-references: EMBL:Z50873; PIDN:CAA90762.1; GSPDB:GN00028; CESP:F19H6.1
 A:Experimental source: clone F17E5
 R:McMurray, A.
 submitted to the EMBL Data Library, November 1995
 A:Reference number: Z19378
 A:Accession: T21124
 A:Status: preliminary; translated from GB/EMBL/DBJ
 A:Molecule type: DNA
 A:Residues: 1-294 <WIL>
 A:Cross-references: EMBL:Z68115; PIDN:CAA92169.1; GSPDB:GN00028; CESP:F19H6.1
 A:Experimental source: clone F19H6
 C:Genetics:
 A:Gene: CESP:F19H6.1
 A:Map position: X
 A:Introns: 55/3; 85/2; 113/3; 150/1; 192/3; 266/2

Query Match 11.7%; Score 424; DB 2; Length 294;
 Best Local Similarity 37.1%; Pred. No. 2e-10;
 Matches 85; Conservative 55; Mismatches 83; Indels 6; Gaps 4;

QY 1 MDKYDVIRKAGGAFGKAYLAGKSDSKHCIVKEINFEKMPIOEKEASKKEVILLKMK 59
 Db 20 LEUFIETKIGKGGQSEVFRAQCTWDLHLVALUKIQVEMVDQKARQDCLKIDLLKQIN 79
 QY 60 HPNIVAFNSFQNGRLFTVMEYCDGDLMKRIN--RQRGLVFSBDQILGMFVQISLGLK 117
 Db 80 HYNVIRYASFIDNNQNLNIVLELAGDMSRMKHKFKGRLIPEKTIWKVQVLARALA 139


```

118 HTHDRKILHRDIAQNIIFLSKNGMYAKIIGDGIARVLNNMELARTICIGPYIYLSPIQ 177
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
140 HHMSKIRMRHDIKPANVFTITNGIV-KLGD-LGLGRFFSKTTAAHSLVGTPIYMSPERIQ 198
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
178 NRPYNNKTDIWSLGVLCYIELCTLKHPPFEGN--NLQQLVLKICQAHFAPI 224
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|
199 ESGYNFKSDLWSTGCLLYEMAALQSPFYGDKMNLVSLCKKIENCFEPL 247
   ::::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|::|

RESULT 13
23580
probable protein kinase KIN3 (EC 2.7.1.1-) - yeast (Saccharomyces cerevisiae)
;Alternate names: probable protein kinase FUN52; probable protein kinase NPX1; protein
;Species: Saccharomyces cerevisiae
;Date: 04-Dec-1992 #sequence_revision 04-Dec-1992 #text_change 24-Sep-1999
;Accession: S23580; S20123; S40906; S43451; S11185
;Schweitzer, B.; Philippsen, P
;Mol. Gen. Genet. 234, 164-167, 1992
;Title: NPX1, a nonessential protein kinase gene in Saccharomyces cerevisiae with simil
;Reference number: S23580; MUID:92357012; PMID:1495480
;Accession: S23580
;Molecule type: DNA
;Residues: 1-435 <SCH>
;Cross-references: EMBL:X50549; NID:g298024; PIDN:CAA43042.1; PID:g298026
;Davies, C.J.; Hutchison III, C.A.
;Submitted to the EMBL Data Library, January 1992
;Description: A directed DNA sequencing strategy based upon Tn3 transposon mutagenesis:
;Reference number: S20123
;Accession: S20123
;Molecule type: DNA
;Residues: 1-435 <DAV>
;Cross-references: EMBL:M67445
;Clark, M.W.; Keng, T.; Storms, R.K.; Zhong, W.; Fortin, N.; Zeng, B.; Delaney, S.; Oue
;Submitted to the EMBL Data Library, November 1993
;Description: Sequencing of chromosome I of Saccharomyces cerevisiae: Analysis of the 4
;Reference number: S40891
;Accession: S40906
;Molecule type: DNA
;Residues: 1-430 <CLA>
;Cross-references: EMBL:L22015; MIPS:YAR018C
;Clark, M.W.; Keng, T.; Storms, R.K.; Zhong, W.; Fortin, N.; Zeng, B.; Delaney, S.; Oue
;Submitted to the EMBL Data Library, November 1993
;Title: Sequencing of chromosome I of Saccharomyces cerevisiae: analysis of the 42 kbp
;Reference number: S43441; MUID:95028152; PMID:7941740
;Accession: S43451
;Status: nucleic acid sequence not shown; translation not shown
;Molecule type: DNA
;Residues: 1-435 <CLW>
;Cross-references: EMBL:L22015; NID:gl339990; PIDN:AA04964.1; PID:g349756
;Note: The nucleotide sequence was submitted to the EMBL Data Library, November 1993
;Jones, D.G.L.; Rosamond, J.
;Gene 90, 87-92, 1990
;Title: Isolation of a novel protein kinase-encoding gene from yeast by oligodeoxyribon
;Reference number: S11185; MUID:90337351; PMID:2199332
;Accession: S11185
;Status: not compared with conceptual translation
;Molecule type: DNA
;Residues: 1-97, 'L', '99-120, 'Q', '122-129, 'I', '131-234, 'SQIS', '239-357, 'R', '359-369 <JON>
;Cross-references: EMBL:M55416
;Experimental source: strain SB303
;Genetics:
;Gene: SGD:KIN3; NPX1; FUN52
;Cross-references: SGD:S0000071; MIPS:YAR018C
;Map position: 1R
;Superfamily: unassigned Ser/Thr or Tyr-specific protein kinases; protein kinase homolo
;Keywords: ATP; phosphotransferase; serine/threonine-specific protein kinase
;31-39/Region: protein kinase ATP-binding motif
;Query Match 11.7%; Score 423; DB 2; Length 435;
;Best Local Similarity 27.4%; Fred. No. 3.1e-10;
;Matches 123; Conservative 69; Mismatches 141; Indels 116; Gaps 14;

```

```

Qy 3 KYDVITAKGQGAFAKAYIAKQSKCHVCVKEINFEKMPIQEKEAKKVEILLERKWHFN 62
Db 24 EYQVLEEETGRSGVSGVRVYIHPTKLLVARDIKYGHMNSKERQQULIACSLSSQLKHEN 83
Qy 63 IVAFFN-SFOENGR-LFTVMVYCDGGDLMKRIN--RQGVLFSEDOIGLWFWQVLSLGLKH 118
Db 84 IVEFYNWDFDEQKEVLYLYMYEYCSRGDLSQWKHYQKEHYIPEKIVWGLLAQLLTALYK 143
Qy 119 IH-----DRK-----ILHDIKAQNIIFS----- 137
Db 144 CHYGVELPLTATTYIDRMKPPVKGNIVIHRLDKPGNIFLSYDDSDYNINEQVDGHEEVNS 203
Qy 138 -----KNG-----MVAKLGDPGIARVLNNSMELAPTCIGTVPYLSPEICQNK 179
Db 204 NYYDRHVNNSKRGSPMDYSQVVVYKLGDFGLAKSLETSTQFATTYGTYFYNSPEVLMQD 263
Qy 180 PYNKTDIWSJGCVLYELCTLKHPFEGNNIQQVLKICQAHFAPISPGFSRELHLSLSOL 239
Db 264 FYSPLSDIWSJGCVIFEMCSLHPFPQAKNYLEIQTKIKGCKDTPVEYYSRGLNAIHSM 323
Qy 240 FOVSFRDRPSINSILKRPFLNLPKYLTPVEIQEFSFHMILCRAGAPASRHAGKV--- 296
Db 324 IDVNLRTPS-----TELIQD--IQIRTARKSLQLEPERKLLDYE 363
Qy 297 -QCKIKIQKVRFGKCPPRSIRSVPIKKNAILHNENR-----PPAGAQAARSIKMIERP 349
Db 364 NELTNIEMI-----LEKQAIYEYRELSQLKEQFTQAVEERAREV--ISGK 406
Qy 350 KIAACVGHYDYYAQLDMRRERAKPSYH 378
Db 407 KYGVPEISINGYG-----KKFAKPAYH 429

RESULT 14
B96587
hypothetical protein F20D21.33 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 31-Mar-2001
C:Accession: B96587
R:Theologos, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.;
  Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; I
  ansen, N.P.; Hughes, B.; Huizar, L.
  Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.
  C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.;
  Rizzo, M.; Rozenberg, T.; Rowley, D.; Sakano, H.
  A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H. H.
  ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
  A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
  A:Reference number: A96141; MUID:21016719; PMID:11130712
A:Accession: B96587
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-200 <STO>
A:Cross-references: GB:AB005173; NID:g4585993; PID:AAD52629.1; GSPDB:GN00141
C:Genetics:
A:Gene: F20D21.33
A:Map position: 1

```

[illegible]

[illegible]